

Protein Sequence Searches - 10/8/04

All of the sequence databases on the ABSS have been updated. A change has occurred in the protein databases.

- Two protein databases, SPTREMBL and SwissProt, are now produced as a single, merged database called UniProt.
- Results from UniProt have the file extension **.rup**.
- Sequences in UniProt are identified by the same ID that had been used in SPTREMBL or SwissProt.
- In instances where the database curators have determined that an SPTREMBL record and a SwissProt record represent the same sequence, the two records have been merged into one. Both IDs are present in the record. Any differences found between the two sequences are recorded in the FT (feature table) fields.

If you have any questions regarding these changes or your results, please contact any STIC searcher.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 18:22:47 ; Search time 196 Seconds
(without alignments)
1441.372 Million cell updates/sec

Title: US-10-017-724-6

Perfect score: 2676

Sequence: 1 MKRMVSWSHKLTMKHLL.....GSWYSMRKVMKIRPFPOQ 491

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 segs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2671	99.8	491	1 FIBB HUMAN	P02675 homo sapien
2	2259	84.4	481	2 Q8K0E8	Q8K0E8 mus musculus
3	2176	81.3	479	1 FIBB RAT	P14480 rattus norv
4	2135	79.8	468	1 FIBB BOVIN	P02676 bos taurus
5	2080	77.7	503	2 Q7TME5	Q7TME5 rattus norv
6	1744	65.2	463	1 FIBB CHICK	Q02020 gallus gall
7	1734.5	64.8	488	2 Q91589	Q91589 xenopus lae
8	1464	54.7	485	2 Q6NVE1	Q6NVE1 brachydanio
9	1464	54.7	485	2 A4H66629	A4H66629 brachydan
10	1384.5	51.7	321	2 Q7SYN8	Q7SYN8 xenopus lae
11	1279	47.8	236	2 Q12P1	Q12P1 mus musculus
12	1211	45.3	477	1 FIBB PETMA	P02678 petromyzon
13	657	24.6	431	2 Q7ZVGT	Q7ZVGT brachydanio
14	656.5	24.5	435	2 Q93568	Q93568 gallus gall
15	631.5	23.6	436	2 Q8VCM7	Q8VCM7 mus musculus
16	630.5	23.6	437	2 A4P35744	A4P35744 homo sapi
17	630.5	23.6	453	1 FIBB HUMAN	P02679 homo sapien
18	611	22.8	439	2 Q7SZ13	Q7SZ13 xenopus lae
19	609.5	22.8	399	2 Q7Z664	Q7Z664 homo sapien
20	609	22.8	444	1 FIBG BOVIN	P12799 bos taurus
21	607.5	22.7	438	1 FIBG XENLA	P17634 xenopus lae
22	601.5	22.5	445	1 FIBG RAT	P02680 rattus norv
23	598.5	22.4	432	1 FIBG PETMA	Q4115 petromyzon
24	575	21.5	316	2 Q28529	Q28529 mustela put
25	563	21.0	313	2 Q6GILL5	Q6GILL5 xenopus lae
26	545	20.4	760	2 Q7ZTR1	Q7ZTR1 xenopus lae
27	544.5	20.3	866	1 FIBA HUMAN	P02671 homo sapien
28	535	20.0	764	2 Q6GNK6	Q6GNK6 xenopus lae
29	534.5	20.0	235	2 Q28763	Q28763 papio cynoc
30	531.5	19.9	493	1 ANL2 MOUSE	Q9R045 mus musculus
31	531.5	19.9	493	2 Q8BM09	Q8BM09 mus musculus

32	530.5	19.8	493	1 ANL2 HUMAN	Q9UKU9 homo sapien
33	529.5	19.8	493	2 Q9JU03	Q9JJ03 rattus norv
34	523.5	19.6	782	2 Q7TQ70	Q7TQ70 rattus norv
35	522	19.5	314	2 Q71KU9	Q71KU9 mus musculus
36	522	19.5	314	2 AAQ05798	AAQ05798 mus muscu
37	517	19.3	314	2 Q8VC25	Q8VC25 mus musculus
38	516.5	19.3	741	1 FIBA CHICK	P14448 gallus gall
39	512.5	19.2	782	1 FIBA RAT	P06399 rattus norv
40	512	19.1	312	1 FGLI HUMAN	Q08830 homo sapien
41	512	19.1	312	2 Q8NG32	Q8NG32 homo sapien
42	512	19.1	312	2 AAP35281	AAp35281 homo sapi
43	510.5	19.1	491	2 Q95841	Q95841 homo sapien
44	510.5	19.1	491	2 AAQ88645	AAq88645 homo sapi
45	505	18.9	236	2 Q8NZJ9	Q8NZJ9 homo sapien

ALIGNMENTS

RESULT 1

ID	FIBB_HUMAN	STANDARD;	PRT;	491 AA.
AC	P02675;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Fibrinogen beta chain precursor [Contains: Fibrinopeptide B].			
GN	Name=FGB;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91344740; PubMed=2102623;			
RA	Chung D.W., Harris J.E., Davie E.W.;			
RT	"Nucleotide sequences of the three genes coding for human			
RT	fibrinogen.";			
RL	Adv. Exp. Med. Biol. 281:39-48(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83283433; PubMed=6688356;			
RA	Chung D.W., Que B.G., Rixon M.W., Mace M. Jr., Davie E.W.;			
RT	"Characterization of complementary deoxyribonucleic acid and genomic			
RT	deoxyribonucleic acid for the beta chain of human fibrinogen.";			
RL	Biochemistry 22:3244-3250(1983).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Chung D.W., Harris J.E., Davie E.W.;			
RT	"Nucleotide sequences of the three genes coding for human			
RT	fibrinogen.";			
RL	(in) Liu C.Y., Chien S. (eds.);			
RL	Fibrinogen, thrombosis, coagulation and fibrinolysis, pp.39-48, Plenum			
RL	Press, New York (1991).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANTS SER-100; HIS-170; LEU-265 AND			
RP	LYS-478.			
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Foel C.L., Yi Q.,			
RA	Nickerson D.A.;			
RT	"SeattlesNPs. NHLBI HL66682 program for genomic applications, UW-			
RT	FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE OF 31-491, AND CARBOHYDRATE-LINKAGE SITE.			
RA	Henschen A., Lottspeich F., Southan C., Topfer-Petersen E.;			
RT	"Human fibrinogen: sequence, sulfhydryl bridges, glycosylation and some			
RT	structural variants.";			
RL	(in) Peeters H. (eds.);			
RL	Proides of the biological fluids, Proc. 28th colloquium, pp.51-56,			
RL	Pergamon Press, Oxford (1980).			
RN	[6]			
RP	SEQUENCE OF 31-491.			
RX	MEDLINE=79124640; PubMed=420779;			

- RA Watt K.W.K., Takagi T., Doolittle R.F.;
 RT "Amino acid sequence of the beta chain of human fibrinogen.";
 RL Biochemistry 18:68-76(1979).
 RN [7]
- RP SEQUENCE OF 31-148, AND DISULFIDE BONDS.
 RX MEDLINE=76225080; PubMed=936108;
 RA Blomback B., Hessel B., Hoeg D.;
 RT "Disulfide bridges in NH2-terminal part of human fibrinogen.";
 RL Thromb. Res. 8:639-658(1976).
 RN [8]
- RN SEQUENCE OF 1-38 FROM N.A.
 RP MEDLINE=871146483; PubMed=3029722;
 RX Huber P., Dalmon J., Courtois G., Laurent M., Assouline Z.,
 RA Marguerie G.;
 RT "Characterization of the 5'-flanking region for the human fibrinogen
 beta gene.";
 RL Nucleic Acids Res. 15:1615-1625(1987).
 RN [9]
- RN SEQUENCE OF 31-44.
 RP Blomback B., Blomback M., Grondahl N.J., Guthrie C., Hinton M.;
 RA "Studies on fibrinopeptides from primates.";
 RL Acta Chem. Scand. 19:1788-1789(1965).
 RN [10]
- RN REVIEW, AND DISULFIDE BONDS.
 RP MEDLINE=83254370; PubMed=6575689;
 RX Henschen A., Lottspeich F., Kehl M., Southan C.;
 RA "Covalent structure of fibrinogen.";
 RL Ann. N. Y. Acad. Sci. 408:28-43(1983).
 RN [11]
- RN DISULFIDE BONDS.
 RP MEDLINE=77245999; PubMed=891553;
 RX Gaardlund B., Hessel B., Marguerie G., Murano G., Blomback B.;
 RA "Primary structure of human fibrinogen. Characterization of disulfide-
 containing cyanogen-bromide fragments.";
 RL Eur. J. Biochem. 77:595-610(1977).
 RN [12]
- RN DISULFIDE BONDS.
 RP Doolittle R.F., Takagi T., Watt K.W.K., Bouma H. III, Cottrell B.A.,
 RA Cassman K.G., Goldbaum D.M., Doolittle L.R., Friesner S.J.;
 RT "The structures of fibrinogen and fibrin.";
 RL (in) Magnussen S., Ottesen M., Folmann B., Dano K., Neurath H.
 (eds.);
 RL Regulatory proteolytic enzymes and their inhibitors, pp.163-172,
 Pergamon Press, New York (1978).
 RN [13]
- RN REVIEW, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.
 RP MEDLINE=84305751; PubMed=6383194;
 RX Doolittle R.F.;
 RA "Fibrinogen and fibrin.";
 RL Annu. Rev. Biochem. 53:195-229(1984).
 RN [14]
- RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 164-491.
 RP MEDLINE=97472408; PubMed=9333233;
 RX Spraggon G., Everse S.J., Doolittle R.F.;
 RA "Crystal structures of fragment D from human fibrinogen and its
 crosslinked counterpart from fibrin.";
 RL Nature 389:455-462(1997).
 RN [15]
- RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 164-491.
 RP MEDLINE=98292395; PubMed=9628725;
 RX Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;
 RA "Crystal structure of fragment double-D from human fibrin with two
 different bound ligands.";
 RL Biochemistry 37:8637-8642(1998).
 RN [16]
- RN X-RAY CRYSTALLOGRAPHY.
 RP MEDLINE=99175089; PubMed=10074346;
 RX Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;
 RA "Conformational changes in fragments D and double-D from human
 fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide.";
 RL Biochemistry 38:2941-2946(1999).
 RN [17]
- RN INTERACTION WITH FBLN1.
 RP
- RX MEDLINE=95370284; PubMed=7642629;
 RA Tran H., Tanaka A., Litvinovich S.V., Medved L.V., Haudenschild C.C.,
 RA Agraves W.S.;
 RT "The interaction of fibulin-1 with fibrinogen. A potential role in
 hemostasis and thrombosis.";
 RL J. Biol. Chem. 270:19458-19464(1995).
 RN [18]
- RN VARIANT BALTIMORE-2 LYS-478.
 RP MEDLINE=89058942; PubMed=3194892;
 RX Schmeltzer C.H., Ebert R.P., Bell W.R.;
 RA "A polymorphism at B beta 448 of fibrinogen identified during
 structural studies of fibrinogen Baltimore II.";
 RL Thromb. Res. 52:173-177(1988).
 RN [19]
- RN VARIANT ISE ARG-45.
 RP MEDLINE=91208409; PubMed=2018836;
 RX Yoshida N., Wada H., Morita K., Hirata H., Matsuda M., Yamazumi K.,
 RA Asakura S., Shirakawa S.;
 RT "A new congenital abnormal fibrinogen Ise characterized by the
 replacement of B beta glycine-15 by cysteine.";
 RL Blood 77:1958-1963(1991).
 RN [20]
- RN VARIANT NAPLES THR-98.
 RP MEDLINE=92340664; PubMed=1634610;
 RX Koopman J., Haverkate F., Lord S.T., Grimbergen J., Mannucci P.M.;
 RA "Molecular basis of fibrinogen Naples associated with defective
 thrombin binding and thrombophilia. Homozygous substitution of B beta
 68 Ala-->Thr.";
 RL J. Clin. Invest. 90:238-244(1992).
 RN [21]
- RN VARIANTS IJMUUDEN CYS-44 AND NIJMEGEN CYS-74.
 RP MEDLINE=92228809; PubMed=1565641;
 RX Koopman J., Haverkate F., Grimbergen J., Engesser L., Novakova I.,
 RA Kerst A.F.J.A., Lord S.T.;
 RT "Abnormal fibrinogens IJmuuden (B beta Arg14-->Cys) and Nijmegen (B
 beta Arg44-->Cys) form disulfide-linked fibrinogen-albumin
 complexes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:3478-3482(1992).
 RN [22]
- RN VARIANT NEW YORK-1 39-GLY--LEU-102 DEL.
 RP MEDLINE=85157605; PubMed=3156856;
 RX Liu C.Y., Koehn J.A., Morgan F.J.;
 RA "Characterization of fibrinogen New York 1. A dysfunctional fibrinogen
 with a deletion of B beta(9-72) corresponding exactly to exon 2 of the
 gene.";
 RL J. Biol. Chem. 260:4390-4396(1985).
 RN [23]
- RN VARIANTS GLU-2; LEU-265 AND LYS-478.
 RP MEDLINE=99318093; PubMed=10391209; DOI=10.1038/10290;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [24]
- RN ERRATUM.
 RP Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 RN [25]
- RN VARIANTS CONGENITAL AFIBRINOGENEMIA ARG-383 AND ASP-430.
 RP MEDLINE=20129589; PubMed=10666208;
 RX Duga S., Asselta R., Santagostino E., Zeinali S., Simonc T.,
 Maicovati M., Mannucci P.M., Tencini M.L.;
 RT "Missense mutations in the human beta fibrinogen gene cause congenital
 afibrinogenemia by impairing fibrinogen secretion.";
 RL Blood 95:1336-1341(2000).
 RN [26]
- RN VARIANT CONGENITAL AFIBRINOGENEMIA CYS-196.
 RP

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 18:25:22 ; Search time 42 Seconds

(without alignments)
1124,820 Million cell updates/sec

Title: US-10-017-724-6

Perfect score: 2676

Sequence: 1 MKRWVSWSFHLKTKMKHLL.....GWSYMKRWKMKIRPFPPQ 491

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2671	99.8	491	1 FGHUB	fibrinogen beta ch
2	2132	79.7	468	1 FGHUB	fibrinogen beta ch
3	1744	65.2	463	2 A38463	fibrinogen beta ch
4	1546	57.8	328	2 A05299	fibrinogen beta ch
5	1218	45.5	479	2 A25052	fibrinogen beta ch
6	630.5	23.6	437	1 FGHUG	fibrinogen gamma-A
7	630.5	23.6	453	1 FGHUGB	fibrinogen gamma-A
8	609	22.8	444	2 S05313	fibrinogen gamma-B
9	607.5	22.7	438	2 A32670	fibrinogen gamma c
10	598.5	22.4	432	1 FGLMGS	fibrinogen gamma c
11	590.5	22.1	437	1 FGRTGA	fibrinogen gamma-A
12	590.5	22.1	445	1 FGRTGB	fibrinogen gamma-B
13	544.5	20.3	866	2 D4234	fibrinogen alpha c
14	529.5	19.8	493	2 JC8027	type 1 angiotensin
15	532	19.1	312	2 JN0596	fibrinogen-related
16	496	18.5	641	1 A41932	fibrinogen alpha-I
17	464	17.3	432	2 A27447	cytotoxic T-lympho
18	461	17.2	439	2 I37391	fibrinogen-like pr
19	460	17.2	432	2 I56934	fibrinogen-like pr
20	437	16.3	334	2 JC5980	ficollin-A precursor
21	430.5	16.1	326	2 S61517	ficollin-1 precursor
22	429.5	16.1	333	2 A47172	transforming growt
23	426.5	15.9	326	2 B47172	ficollin-beta - pig
24	422.5	15.8	1353	1 JH0675	restrictin precursor
25	414	15.5	1356	2 A45445	janusin precursor
26	411.5	15.4	2019	1 JQ1322	tenascin precursor
27	406.5	15.2	1810	1 A32230	tenascin precursor
28	401	15.0	91	2 S69115	fibrinogen beta ch
29	395.5	14.8	1746	1 S19694	tenascin precursor

ALIGNMENTS

RESULT 1

FGHUB

fibrinogen beta chain precursor [validated] - human

N;Alternate names: coagulation factor I

C;Contains: fibrinopeptide B

C;Species: Homo sapiens (man)

C;Date: 24-Apr-1984 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C;Accession: B43568; A90469; I37389; A94433; A90437; A94309; G54223; A03121; B371

R;Chung, D.W.; Harris, J.E.; Davie, E.W.

Adv. Exp. Med. Biol. 281, 39-48, 1990

A;Title: Nucleotide sequences of the three genes coding for human fibrinogen.

A;Reference number: A43568; MUID:91344740; PMID:2102623

A;Accession: B43568

A;Molecule type: DNA

A;Residues: 9-191, 'P', 193-491 <CHU>

A;Cross-references: UNIPROT:P02675

R;Chung, D.W.; Que, B.G.; Rixon, M.W.; Mace Jr., M.; Davie, E.W.

Biochemistry 22, 3244-3250, 1983

A;Title: Characterization of complementary deoxyribonucleic acid and genomic deoxyribonuc

A;Reference number: A90469; MUID:83283433; PMID:6688356

A;Accession: A90469

A;Molecule type: DNA

A;Residues: 1-38 <CH1>

A;Accession: B90469

A;Molecule type: mRNA

A;Residues: 9-191, 'A', 193-491 <CH2>

A;Cross-references: GB:J00129; NID:g182429; PIDN:AAA52429.1; PID:g182430

R;Huber, P.; Dalmon, J.; Courtois, G.; Laurent, M.; Assouline, Z.; Marguerie, G.

Nucleic Acids Res. 15, 1615-1625, 1987

A;Title: Characterization of the 5'-flanking region for the human fibrinogen beta gene.

A;Reference number: I37389; MUID:87146483; PMID:3029722

A;Accession: I37389

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-38 <HUB>

A;Cross-references: EMBL:X05018; NID:g31400; PIDN:CAA28674.1; PID:g31401

R;Henschen, A.; Lottspeich, F.; Southan, C.; Topfer-Petersen, E.

In Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Per

A;Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural va

A;Reference number: A94433

A;Contents: carbohydrate binding

A;Accession: A94433

A;Molecule type: protein

A;Residues: 31-137, 'QS', 140-144, 'QF', 147-491 <HEN>

R;Watt, K.W.K.; Takagi, T.; Doolittle, R.F.

Biochemistry 18, 68-76, 1979

A;Title: Amino acid sequence of the beta chain of human fibrinogen.

A;Reference number: A90437; MUID:79124640; PMID:420779

A;Accession: A90437

A;Molecule type: protein

A;Residues: 31-144, 'QF', 147-231, 'D', 233-330, 'E', 332-491 <WAT>

R;Blomback, B.; Hessel, B.; Hogg, D.

tenascin Y precurs
tenascin-X precurs
tenascin homolog -
tenascin-X - pig (
tenascin-X - mouse
probable tenascin
tenascin-X - bovin
fibrinogen-related
tenascin-C - human
beta-fibrinogen -
scabrous locus (sc
hypothetical prote
hypothetical prote
tenascin - African
hypothetical prote
hypothetical prote

30 394.5 14.7 1914 2 T42635
31 390.5 14.6 3586 1 A40701
32 389.5 14.6 220 2 S28170
33 389.5 14.6 417 2 S65944
34 386.5 14.4 860 2 I48839
35 386.5 14.4 4006 2 T09070
36 385.5 14.4 4135 2 T42629
37 378.5 14.1 282 2 A35084
38 365 13.6 2201 2 A32160
39 334 12.5 55 2 I67595
40 329.5 12.3 774 2 A39832
41 223 8.3 463 2 T15876
42 201 7.5 431 2 T29850
43 173.5 6.5 77 2 I51647
44 169.5 6.3 915 2 T21773
45 169.5 6.3 927 2 T21772

Thromb. Res. 8, 639-658, 1976
 A;Title: Disulfide bridges in NH-2-terminal part of human fibrinogen.
 A;Reference number: A94309; PMID:76225080; PMID:936108
 A;Contents: disulfide bonds
 A;Accession: A94309
 A;Molecule type: protein
 A;Residues: 31-112, 'E', 114-137, 'QS', 140-144, 'OF', 147-148 <BLO>
 R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.
 Biochemistry 33, 1988-1992, 1994
 A;Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins
 A;Reference number: A54223; PMID:94162201; PMID:8117655
 A;Accession: G54223
 A;Molecule type: protein
 A;Residues: 164-174 <KUN>
 R;Henschen, A.; Lottspeich, F.; Kehl, M.; Southan, C.
 Ann. N. Y. Acad. Sci. 408, 28-43, 1983
 A;Title: Covalent structure of fibrinogen.
 A;Reference number: A90037; PMID:83254370; PMID:6575689
 A;Contents: annotation; review; disulfide bonds
 R;Gardlund, B.; Hessel, B.; Marguerie, G.; Murano, G.; Blombach, B.
 Eur. J. Biochem. 77, 595-610, 1977
 A;Title: Primary structure of human fibrinogen. Characterization of disulfide-containing
 A;Reference number: A91249; PMID:77245999; PMID:891553
 A;Contents: annotation; disulfide bonds
 R;Doollittle, R.F.; Takagi, T.; Watt, K.; Bouma III, H.; Cottrell, B.A.; Cassman, K.G.; G
 in Regulatory Proteolytic Enzymes and Their Inhibitors, Magnusson, S., Ottesen, M., Fold
 A;Title: The structures of fibrinogen and fibrin.
 A;Reference number: A94437
 A;Contents: annotation; disulfide bonds
 R;Doollittle, R.F.
 Annu. Rev. Biochem. 53, 195-229, 1984
 A;Title: Fibrinogen and fibrin.
 A;Reference number: A90041; PMID:84305751; PMID:6383194
 A;Contents: annotation; review, EM structure, polymerization, ligands
 R;Chung, D.W.; Rixon, M.W.; Que, B.G.; Davie, E.W.
 Ann. N. Y. Acad. Sci. 408, 449-456, 1983
 A;Title: Cloning of fibrinogen genes and their cDNA.
 A;Reference number: A90038; PMID:83254384; PMID:6575700
 A;Contents: annotation
 R;Kirschbaum, N.B.; Budzynski, A.Z.
 J. Biol. Chem. 265, 13669-13676, 1990
 A;Title: A unique proteolytic fragment of human fibrinogen containing the Aalpha COOH-ter
 A;Reference number: A37117; PMID:90337977; PMID:2143188
 A;Contents: annotation; hementin cleavage site
 A;Note: hementin, a protease from *Haemeteria ghilianii*, the giant South American leech,
 C;Comment: The conversion of fibrinogen to fibrin is triggered by thrombin, which cleave
 ization sites responsible for the formation of the soft clot.
 C;Comment: The soft clot is converted into the hard clot by factor XIIIa (fibrin-stabili
 ger) and between alpha chains (weaker) of different monomers.
 C;Comment: All fibrinogen chains are synthesized in the liver.
 C;Genetics:
 A;Gene: GDB:FGB
 A;Cross-references: GDB:119130; OMIM:134830
 A;Map position: 4q28-4q28
 A;Introns: 38/3; 102/3; 164/1; 240/1; 278/1; 320/1; 415/2
 C;Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR:FC
 ins are contained in the core. Two three-chain coiled coils emerge from this core and co
 from the distal domain nodes.
 C;Function:
 A;Description: fibrinogen cleaved by thrombin yields monomers that are polymerized into
 A;Pathway: blood coagulation
 C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfid
 C;Keywords: blood coagulation; coiled coil; glycoprotein; liver; plasma; pyroglyutamic ac
 F;1-30/Domain: (or 4-30 or 15-30) signal sequence #status predicted <SIG>
 F;31-491/Product: fibrinogen beta chain #status experimental <MAT>
 F;31-44/Product: fibrinopeptide B #status experimental <APT>
 F;45-491/Product: fibrin beta chain #status experimental <PGB>
 F;45-47/Region: polymerization site
 F;99-228/Domain: fibrinogen disulfide ring homology <PDR>
 F;238-487/Domain: fibrinogen beta/gamma homology <FBG>
 F;31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F;44-45/Cleavage site: Arg-Gly (thrombin) #status experimental

F;95/Disulfide bonds: interchain (to alpha-55) #status experimental
 F;106/Disulfide bonds: interchain (to alpha-68) #status experimental
 F;110/Disulfide bonds: interchain (to gamma-45) #status experimental
 F;223/Disulfide bonds: interchain (to alpha-184) #status experimental
 F;227/Disulfide bonds: interchain (to gamma-161) #status experimental
 F;231-316,241-270,424-437/Disulfide bonds: #status experimental
 F;394/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 99.8%; Score 2671; DB 1; Length 491;
 Best Local Similarity 99.8%; Pred. No. 3.1e-170;
 Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MKRVSWSFHKLKTKMKHLLLLCVFLVKSQGVNDNEEGFFSARGHRPLDKREAPSLR	60
DB	1	MKRVSWSFHKLKTKMKHLLLLCVFLVKSQGVNDNEEGFFSARGHRPLDKREAPSLR	60
QY	61	PAPPIISGGYRPAKAAATOKKVERKAPDAGGCHADPDGLVLCPTGCOEALLQOE	120
DB	61	PAPPIISGGYRPAKAAATOKKVERKAPDAGGCHADPDGLVLCPTGCOEALLQOE	120
QY	121	RPTRNSVDELNNVEAVSQTSSSFQYMYLLKDLWKQKQVKNENNVNYSSELEKHQ	180
DB	121	RPTRNSVDELNNVEAVSQTSSSFQYMYLLKDLWKQKQVKNENNVNYSSELEKHQ	180
QY	181	LYIDTVNSNTATNLRVLSILENLSKIQLKESDVSAQMEYCRTPCTVSCNIPVVS	240
DB	181	LYIDTVNSNTATNLRVLSILENLSKIQLKESDVSAQMEYCRTPCTVSCNIPVVS	240
QY	241	CEEIRKGGTSEMVLIQDSSVKPYRYCDMNTENGWTVIONRODGSVDGRKWD	300
DB	241	CEEIRKGGTSEMVLIQDSSVKPYRYCDMNTENGWTVIONRODGSVDGRKWD	300
QY	301	QGFNVATVTDGKNCGLPGEYWLGNDKISQLTRMGPTLELLEMDKGDVKYKAHYGG	360
DB	301	QGFNVATVTDGKNCGLPGEYWLGNDKISQLTRMGPTLELLEMDKGDVKYKAHYGG	360
QY	361	VQNEANKYQISVNVKRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDND	420
DB	361	VQNEANKYQISVNVKRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDND	420
QY	421	RKQCKEDGGGWWYRCHAAENGRYWGQVYTDMAKHGTDGTVVMWVWVWVWVWVW	480
DB	421	RKQCKEDGGGWWYRCHAAENGRYWGQVYTDMAKHGTDGTVVMWVWVWVWVW	480
QY	481	SMKIRPFFPQQ	491
DB	481	SMKIRPFFPQQ	491

RESULT 2

FGBOB

fibrinogen beta chain - bovine
 N;Contains: fibrinopeptide B
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 29-Jul-1981 #sequence revision 29-Jul-1981 #text_change 13-Sep-1996
 C;Accession: A03122; B03117; B37507; A37513; S02443
 R;Blombach, B.; Doollittle, R.F.
 Acta Chem. Scand. 17, 1816-1819, 1963
 A;Title: The sequence of amino acids at the N-terminal end of bovine fibrinopeptide B.
 A;Reference number: A03122
 A;Accession: A03122
 A;Molecule type: protein
 A;Residues: 1-4 <BLO>
 R;Sjoquist, J.; Blombach, B.; Wallen, P.
 Ark. Kemi 16, 425-436, 1960
 A;Title: Amino acid sequence of bovine fibrinopeptides.
 A;Reference number: A03117
 A;Accession: B03117
 A;Molecule type: protein
 A;Residues: 5-21 <SJO>
 R;Martinelli, R.A.; Inglis, A.S.; Rubira, M.R.; Hageman, T.C.; Hurrell, J.G.R.; Leach, S.
 Arch. Biochem. Biophys. 192, 27-32, 1979
 A;Title: Amino acid sequences of portions of the alpha and beta chains of bovine fibrin

Gene 75, 131-138, 1989

C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide
C:Keywords: blood coagulation; glycoprotein; liver; plasma
F:33-65/Domain: fibrinogen disulfide ring homology (fragment) <FDR>
F:75-324/Domain: fibrinogen beta/gamma homology <FBG>
F:231/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 57.8%; Score 1546; DB 2; Length 328;
Best Local Similarity 85.6%; Pred. No. 1.4e-95;
Matches 273; Conservative 22; Mismatches 24; Indels 0; Gaps 0;

QY 173 SSELKHLQYIDETVNSNIATNLRLVLSILNLSKIQKLESQVSAQMEYCRTPCTVSCN 232
DB 10 SVSLVOTQAATDSKVDILSARRVLRSLILEDLRSKMQKLESQVSAQMEYCRTPCTVSCN 69
QY 233 IPVSGKCEBIIIRKGETSEMYLQPPSSVKKPYRYVCDMKTENGQVTVIQRQDGVDF 292
DB 70 IPVSGKCEBIIIRKGETSEMYLQPPSSVKKPYRYVCDMKTENGQVTVIQRQDGVDF 129
QY 293 GRKWDPPYKGFQNVATNTDGNKYCCLPGEYMLGNDKISQLTRMGPTTELLIEMEDWKGDKV 352
DB 130 GRKWDPPYKGFQNVATNTDGNKYCCLPGEYMLGNDKISQLTRMGPTTELLIEMEDWKGDKV 189
QY 353 KAHYGGFTVQNEANKYQISVNVKRYGTAGNALMDGASQALMGENTRTWIHNGMPPFTYDRDN 412
DB 190 KAHYGGFTVQNEANKYQISVNVKRYGTAGNALMDGASQALMGENTRTWIHNGMPPFTYDRDN 249
QY 413 DGLTSDPRKCKSKEDGGWYVNRCHAANPNRGYYWGGYTWDMAKHGTDDGVVWNNWKG 472
DB 250 DGLTSDPRKCKSKEDGGWYVNRCHAANPNRGYYWGGYTWDMAKHGTDDGVVWNNWKG 309
QY 473 SWYSMRKMSKIRPFPQQ 491
DB 310 SWYSMRKMSKIRPFPQQ 328

RESULT 5
A:25052
N:Contains: fibrinogen beta chain - sea lamprey (fragments)
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 25-Oct-1987 #sequence revision 19-Feb-1999 #text_change 09-Jul-2004
R:Bohonus, V.L.; Doolittle, R.F.; Pontes, M.; Strong, D.D.
Biochemistry 25, 6512-6516, 1986
A:Title: Complementary DNA sequence of lamprey fibrinogen beta chain.
A:Reference number: A25052; MUID:87076582; PMID:3790537
A:Accession: A25052
A:Molecule type: mRNA
A:Residues: 39-479 <BOH>
A:Cross-references: UNIPROT:P02678; GB:M14773; NID:G213191; PIDN:AAA49261.1; PID:G213192
R:Cottrell, B.A.; Doolittle, R.F.
Biochim. Biophys. Acta 453, 426-438, 1976
A:Title: Amino acid sequences of lamprey fibrinopeptides A and B and characterization of
A:Reference number: A03120; MUID:177065679; PMID:999898
A:Accession: A03124
A:Molecule type: protein
A:Residues: 1-36 <COT1>
A:Accession: B03124
A:Molecule type: protein
A:Residues: 37-42 <COT2>
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide
C:Keywords: blood coagulation; glycoprotein; sulfoprotein
F:1-36/Product: fibrinopeptide B #status experimental <FPB>
F:37-479/Product: fibrin beta chain #status experimental <MAT>
F:90-219/Domain: fibrinogen disulfide ring homology <FDR>
F:229-477/Domain: fibrinogen beta/gamma homology <FBG>
F:13/Binding site: sulfate (Tyr) (covalent) #status experimental
F:27/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 45.5%; Score 1218; DB 2; Length 479;
Best Local Similarity 48.4%; Pred. No. 1.7e-73;
Matches 223; Conservative 75; Mismatches 151; Indels 12; Gaps 5;

QY 34 NDNEEGFSARHGRPLDK-KREAEPSLRPAPPPISGGYRARPAPAKAAATQKKVEKAPDA 92
DB 26 SNNTAAALDVRGVRPLPSGTRVRRLRHR--RLAPGAVMSRDPSPAPPOEAKAIRDE 83
QY 93 GGCULHAPDGLVLCPTGCGQLQALQOERPRNSVDELNNN-----VEAVSQTSSSSQYM 148
DB 84 GGCMLPESDLGLVLCPTGCGELRELLKQRPVRYKISMLKQNLTYFINSFDRMADS---- 139
QY 149 YLLKDLQKROKQKDNENVVNSYSELEKHQYIDETVNSNIATNLRLVLSILNLSK 208
DB 140 NTLKQNVOTLRRRLNSRSSSTHVNAKEIENRYKVKRIESTVAGSLRSMKSVLEHLRAK 199
QY 209 IQKLESQVSAQMEYCRTPCTVSCNIPVSGKCEBIIIRKGETSEMYLQPPSSVKKPYRYV 268
DB 200 QRMEEAIKTQKELCSAPCTVNCRVVSGMHCEDIYRNGGRTSEAYIQPDLFSEPKV 259
QY 269 YCDMNTENGQVTVIQRQDGVDFGRKWDPPYKGFQNVATNTDGNKYCCLPGEYMLGNDK 328
DB 260 FCDMESHGGGWTVVQNRVDSNPFARDMNTYKAEFGNIAFG-NGKSI CNIPGEYWLGTKT 318
QY 329 ISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQNEANKYQISVNVKRYGTAGNALMDGAS 388
DB 319 VHQLTKQHTQOVLFDMSDWEGSSVYVAQYASRPENEAQGYRLWVEDYSGNAGNALLEGAT 378
QY 389 QLMGENTRTWIHNGMPPFTYDRDNMDGLTSDPRKCKSKEDGGWYVNRCHAANPNRGYYW 448
DB 379 QLMGENTRTWIHNGMPPFTYDRDNMDGLTSDPRKCKSKEDGGWYVNRCHAANPNRGYYW 438
QY 449 GGQYTWDMAKHGTDDGVVWNNWKGSWYSMRKMSKIRPFP 489
DB 439 GGQYTWDMAKHGTDDGVVWNNWKGSWYSMRKMSKIRPFP 479

RESULT 6
FCHUG
N:Alternate names: coagulation factor 1
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence revision 25-Feb-1985 #text_change 09-Jul-2004
A:Accession: A90470; B90494; C94433; B93956; B92448; I37393; A40698; H54223; A03125; C371
R:Chung, D.W.; Chan, W.Y.; Davie, E.W.
Biochemistry 22, 3250-3256, 1983
A:Title: Characterization of a complementary deoxyribonucleic acid coding for the gamma c
A:Reference number: A90470; MUID:83283434; PMID:6688357
A:Accession: A90470
A:Molecule type: mRNA
A:Residues: 1-437 <CHU>
A:Cross-references: UNIPROT:P02679; UNIPROT:Q9UCF3
R:Rixson, M.W.; Chung, D.W.; Davie, E.W.
Biochemistry 24, 2077-2086, 1985
A:Title: Nucleotide sequence of the gene for the gamma chain of human fibrinogen.
A:Reference number: A90494; MUID:85252774; PMID:2990550
A:Accession: B90494
A:Molecule type: DNA
A:Residues: 1-113, '1', 115-437 <RIX>
A:Cross-references: GB:X02415; GB:M10014; NID:G182438; PIDN:AAB59531.1; PID:G182439
R:Henschen, A.; Lottspeich, F.; Southan, C.; Topper-Petersen, E.
in Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Pe
A:Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural va
A:Reference number: A94433
A:Accession: C94433
A:Molecule type: protein
A:Residues: 27-437 <HEN>
R:Kant, J.A.; Lord, S.T.; Crabtree, G.R.
Proc. Natl. Acad. Sci. U.S.A. 80, 3953-3957, 1983
A:Title: Partial mRNA sequences for human Aalpha, Bbeta, and gamma fibrinogen chains: ev
A:Reference number: A93956; MUID:83247396; PMID:6575389
A:Accession: B93956
A:Molecule type: mRNA
A:Residues: 276-437 <KAN>
R:Porreca Jr., A.J.; Cummings, D.E.; Comeau, C.M.; Kant, J.A.; Crabtree, G.R.
J. Biol. Chem. 259, 12826-12830, 1984
A:Title: Structure of the human gamma-fibrinogen gene. Alternate mRNA splicing near the

A;Reference number: A92448; MUID:85030379; PMID:6092346
A;Accession: B92448
A;Molecule type: DNA
R;Imam, A.M.A.; Eaton, M.A.W.; Williamson, R.; Humphries, S.
Nucleic Acids Res. 11, 7427-7434, 1983
A;Title: Isolation and characterisation of cDNA clones for the Aalpha- and gamma-chains
A;Reference number: I37393; MUID:84069777; PMID:6689067
A;Accession: I37393
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 209-270 <RES>
A;Cross-references: EMBL:X00086; NID:g31445; PIDN:CAA24944.1; PID:g577055
R;Bertagnoli, M.E.; Beckerle, M.C.
J. Cell Biol. 121, 1329-1342, 1993
A;Title: Evidence for the selective association of a subpopulation of GPIIb-IIIa with the
A;Reference number: A40698; MUID:93286185; PMID:8509453
A;Accession: A40698
A;Molecule type: protein
A;Residues: 27-33, 'XX', 36-41 <BER>
A;Experimental source: Thrombin-activated platelets
A;Note: sequence extracted from NCBI backbone (NCBIPI:133734)
R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.
Biochemistry 33, 1988-1993, 1994
A;Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins
A;Reference number: A54223; MUID:94162201; PMID:8117655
A;Accession: H54223
A;Molecule type: protein
A;Residues: 27-33, 'XX', 36-41 <KUN>
A;Note: identification of tryptic peptides from high-density lipoproteins
R;Henschen, A.; Lottspeich, F.; Kehl, M.; Southan, C.
Ann. N. Y. Acad. Sci. 408, 28-43, 1983
A;Title: Covalent structure of fibrinogen.
A;Reference number: A90037; MUID:83254370; PMID:6575689
R;Doollittle, R.F.; Takagi, T.; Watt, K.; Bouna III, H.; Cottrell, B.A.; Cassman, K.G.; Folb
in Regulatory Proteolytic Enzymes and Their Inhibitors, Magnusson, S., Ortesen, M., Folb
A;Title: The structures of fibrinogen and fibrin.
A;Reference number: A94437
A;Contents: annotation; disulfide bonds
R;Blomback, B.; Hessel, B.; Hogg, D.
Thromb. Res. 8, 639-658, 1976
A;Title: Disulfide bridges in NH-2-terminal part of human fibrinogen.
A;Reference number: A94309; MUID:76225080; PMID:936108
A;Contents: annotation; disulfide bonds
R;Hoepflich, P.D.; Doollittle, R.F.
Biochemistry 22, 2049-2055, 1983
A;Title: Dimeric half-molecules of human fibrinogen are joined through disulfide bonds
A;Reference number: A90467; MUID:83231465; PMID:6860649
A;Contents: annotation; quaternary structure, disulfide bonds
R;Doollittle, R.F.
Annu. Rev. Biochem. 53, 195-229, 1984
A;Title: Fibrinogen and fibrin.
A;Reference number: A90041; MUID:84305751; PMID:6383194
A;Contents: annotation; review, EM structure, polymerization, ligands
R;Horwitz, B.H.; Varadi, A.; Scheraga, H.A.
Proc. Natl. Acad. Sci. U.S.A. 81, 5980-5984, 1984
A;Title: Localization of a fibrin gamma-chain polymerization site within segment Thr-374
A;Reference number: A94006; MUID:85014892; PMID:6592597
A;Contents: annotation; polymerization region
R;Kloczewiak, M.; Timmons, S.; Lukas, T.J.; Hawiger, J.
Biochemistry 23, 1767-1774, 1984
A;Title: Platelet receptor recognition site on human fibrinogen. Synthesis and structure
A;Reference number: A90483; MUID:84203545; PMID:6326808
A;Contents: annotation; platelet aggregation region
R;Flow, E.F.; Srouji, A.H.; Meyer, D.; Marguerie, G.; Ginsberg, M.H.
J. Biol. Chem. 259, 5388-5391, 1984
A;Title: Evidence that three adhesive proteins interact with a common recognition site
A;Reference number: A92477; MUID:84185664; PMID:6325435
A;Contents: annotation; platelet aggregation region
R;Dang, C.V.; Ebert, R.F.; Bell, W.R.
J. Biol. Chem. 260, 9713-9719, 1985
A;Title: Localization of a fibrinogen calcium binding site between gamma-subunit position

A:Reference number: A92549; MUID:85261382; PMID:3160702
A:Contents: annotation; calcium binding region
F:Kirschbaum, N.E.; Budzynski, A.Z.
J. Biol. Chem. 265, 13669-13676, 1990
A:Title: A unique proteolytic fragment of human fibrinogen containing the Aalpha COOH-ter-
A:Reference number: A37117; MUID:90337977; PMID:2143188
A:Contents: annotation; hementin cleavage site
C:Note: hementin, a protease from Haementeria ghilianii, the giant South American leech,
C:Comment: The conversion of fibrinogen to fibrin is triggered by thrombin, which cleaves
ization sites responsible for the formation of the soft clot.
C:Comment: The soft clot is converted into the hard clot by factor XIIIa (fibrin-stabiliz-
ger) and between alpha chains (weaker) of different monomers.
C:Comment: All fibrinogen chains are synthesized in the liver.
C:Comment: The two forms of gamma chain, A and B (see PIR:FGHUGB), arise by alternate spl-
intron, which makes this chain different from the gamma-B chain at positions 434-437 and
C:Genetics:
A:Gene: GDB:FGG
A:Cross-references: GDB:L19132; OMIM:134850
A:Map position: 4q28-4q28
A:Introns: 26/3; 41/3; 103/1; 134/2; 178/1; 222/3; 284/2; 377/1; 433/3
C:Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR:FG
ins are contained in the core. Two three-chain coiled coils emerge from this core and cor-
from the distal domain nodes.
C:Function:
A:Description: fibrinogen cleaved by thrombin yields monomers that are polymerized into f
A:Pathway: blood coagulation
C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
C:Keywords: alternative splicing; blood coagulation; calcium; coiled coil; glycoprotein;
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-437/Product: fibrinogen gamma-A chain #status experimental <MPT>
F:176-415/Domain: fibrinogen beta/gamma homology <FBG>
F:341-355/Domain: calcium binding #status predicted <CAB>
F:400-422/Region: polymerization site, binding to the amino end of the alpha chain of an
F:423-437/Region: platelet aggregation #status predicted
F:34/Disulfide bonds: interchain (to gamma-35) #status experimental
F:35/Disulfide bonds: interchain (to gamma-34) #status experimental
F:45/Disulfide bonds: interchain (to beta-110) #status experimental
F:49/Disulfide bonds: interchain (to alpha-64) #status experimental
F:78/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:161/Disulfide bonds: interchain (to beta-227) #status experimental
F:165/Disulfide bonds: interchain (to alpha-180) #status experimental
F:179-208,352-365/Disulfide bonds: #status experimental
F:432/cross-link: isopeptide (Gln) (interchain to Lys-432 N6-amino) #status experimental
F:432/cross-link: isopeptide (Lys) (interchain to Gln-424) #status experimental

Query Match	23.6%; Score 630.5; DB 1; Length 437;
Best Local Similarity	33.6%; Pred. No. 1.9e-34;
Matches	142; Conservative 71; Mismatches 140; Indels 69; Gaps 15;
QY	95 CLHADPDLGVLPTGC-----OQEALLQERPRTPNSVDLNNVEAV 137 : : : : :
Db	34 CCILDERFGYCPTTCGIADFLTYOTKVDKIQSLIEDILHQ---VENKTSEVKQLIKAI 90 : : : : :
QY	138 SQT---SSSFQWYLLKDLQRQRQKDNENNVANEYSSELEKHQ---LYIDETVMSNI 191 : : : : :
Db	91 QLTYNPDESSKPNIIDAATL--KSRRMLEE----IMKYASILLTHDSSIFYLOEIYNSN- 143 : : : : :
QY	192 ATNLRVLRILENLRSKIOKLEDVAQMRYCTPCTVCSNIPVSGKECEBIIRKGGET 251 : : : : :
Db	144 --NOKIV-----NLKBEVAQLEAQ-----COEPCKDTVOIHDIITKCQCDIANKGAQ 189 : : : : :
QY	252 SEMYLQPDSKVYPYVCMDNTENGWTVIQNRPQGSVDGRKWDPYKOGFGNVA-TWT 310 : : : :
Db	190 SGIFYFKPLKANQQFLVYCEIDSGNGNWTFQKRLDGSDVFQKNWILOYKEGFQHLSPTGT 249 : : : :
QY	311 DGKNYCGLPGEYWLGNDKISQLTRMG--PTELLEIMEDWKDKVKAHYGGFTVQNEANKY 368 : : : :
Db	250 T-----EFWLGNKEIHLISTQSAPVALRVELEDWNGRTSTADYAMFKVGPPEADKY 300 : : : :
QY	369 QTSVNKYRGTAGNALMDGASQLMGE---NRMTIHNGMFESTYDRONDGLWLTSDPRKQCS 425 : : : :
Db	301 RUTYAYFAGDAGDAFDGDF--FGDDPSDKFTTSHNGMQFSTWDNDNDKF-----EGNCA 353 : : : :

Query Match 22.8%; Score 609; DB 2; Length 444;
 Best Local Similarity 33.0%; Pred. No. 5.3e-33;
 Matches 134; Conservative 77; Mismatches 157; Indels 38; Gaps 12;

QY 95 CLHADPDLGVLPTGQLOEALLOQERPIRNSVDELNNVEAVSQTSSSFQMYLLKDL 154
 DB 32 CCILDERFGSCPTTCGADFLNNYQTSVDKOLRTLEGILYQENKTSBARELVKAIQIS 91

QY 155 WOKROKQVND-ENVVNEVSSELEKHLYIDETVNSNIATNLRLVRSI-----LENLR 206
 DB 92 YNDPQSKNNIESATYKNSKMMEEIMKY--ETLISTHETIRFLQEVVNSNSQKTVNLR 149

QY 207 SKIQKLESVSAQMEYCRTPCTVSCNIPVVSQKECEIIRKGETSEMYLIQDSSVKPY 266
 DB 150 DKVVQLEAN-----COEPCQDTVKIHVDYGEDCDQVANKGAKESGLFIIRP-LKAKQF 201

QY 267 RYCDMNTENGWTVIQNRQDGSVDGRKWDVPYKQFGVGNVATNDGKNGYGLPGEYWLGN 326
 DB 202 LVYCEIDGSGNGWTVFQKRLDGLSLDFKNNIQYKEGFGHLSPTGTGNT-----EFWLGN 255

QY 327 DKISOLTRMG--PTELLIEMEDWKGDKVAHYGGFTVQNEANKYQISVSNKY-RGTAGNAL 383
 DB 256 EXHILISTOSSIPYVIRIQLDENDNGRTSTADYASFVKTGENDKYRLTYAYFTGGDAGDA- 314

QY 384 MDGASQLMGE--NRTMTIHNGMFFSTYDRDNDGMLTSDPRKQCKEDGGGWWYNNRCHAA 440
 DB 315 FDGYD--FGDSSDKFPTSHNGMQFSTWSDNDKY-----DGNCAEQVGLGWWNNKCHAG 367

QY 441 NNNGRYWGQYTWDMAKHGTDDGVVMMNWKSGWSYMRKQSMKIRP 486
 DB 368 HLNGVYVYQGTYSKTSPTNGYDNGIITWATWKRWSYMKKTTMKIIP 413

RESULT 9
 A32670
 fibrinogen gamma chain - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 09-Jul-2004
 R:Pastori, R.L.; Moskaitis, J.E.; Smith Jr., L.H.; Schoenberg, D.R.
 Biochemistry 29, 2399-2605, 1990
 A:Title: Estrogen regulation of Xenopus laevis gamma-fibrinogen gene expression.
 A:Reference number: A32670; MUID:90241882; PMID:2334684
 A:Accession: A32670
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-438 <PAS>
 A:Cross-references: UNIPROT:P17634; GB:J02894; NID:G214139; PIDN:AAA49709.1; PID:G214140
 R:Bhattacharya, A.; Shepard, A.R.; Moser, D.R.; Holland, L.J.
 Mol. Cell. Endocrinol. 72, 213-220, 1990
 A:Title: Isolation and characterization of cDNA clones for the gamma subunit of Xenopus
 A:Reference number: I51416; MUID:91146806; PMID:2289632
 A:Accession: I51416
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-58 <BHA>
 A:Cross-references: GB:M35548; NID:G214141; PIDN:AAA03247.1; PID:G214142
 C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
 C:Keywords: blood coagulation
 F:173-413/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 22.7%; Score 607.5; DB 2; Length 438;
 Best Local Similarity 32.2%; Pred. No. 6.5e-33;
 Matches 137; Conservative 74; Mismatches 146; Indels 69; Gaps 14;

QY 90 PDAGGCLHADPDLGVLPTGQLOEALLOQERPIRNSVDELNNVEAVS-QTSSSFQMY 148
 DB 27 PNTDNCILDRFGYECPTTCGTSDFLNRYQENVDTDLQYLENLLTQISNSTSGTIIIVE 86

QY 149 YLL-----KDLWKRQKQVNDENNVNVEVSSELEKHLYIDETV 187
 DB 87 HLIDSGKKPATSPQTAIDPMTQKSKTCMKW-----LTDKNNY-----QYEENILYQEVY 137

QY 188 NSNIATNLRLVLSILENLSKIQKLESVSAQMEYCRTPCTVSCNIPVVSQKECEIIRK 247
 DB 138 SSN-----QNKIFLLKOKIANLELQ-----CQPCRDVTQIOEFTGDCQCEVANK 182

QY 248 GGETSEMYLIQDSSVVKPYRVYCDMNTENGWTVIQNRQDGSVDGRKWDVPYKQFGVNA 307
 DB 183 GARLSGLYIKPLKAKQOFLVYCEIEPFGSANTVQIQRLLDGSVNFHKNVQVREGFGYLS 242

QY 308 TMTDGNKYGCLPGEYWLGNDKISQLTRMG--PTELLIEMEDWKGDKVAHYGGFTVQNEA 365
 DB 243 PND-----KTEFWLGNKXIHLLSTQSTIPYVVRILELDSNQKSTADYSTFRLGSEK 294

QY 366 NKYQISVSNKY-RGTAGNALMDGASQLMGE--NRTMTIHNGMFFSTYDRDNDGMLTSDPR 421
 DB 295 DNYRFTYAYFIQGDAGDA--FDGFD--FGDDPSDKFYTSHNGMQFSTFDKNDKF-----D 346

QY 422 KQCKEDGGGWWYNNRCHAAANPNRGYVWGQYTWDMAKHGTDDGVVMMNWKSGWSYMRKQ 480
 DB 347 GNCAEQDGGGWWNNRCHAAHLNGKYQGTYSADSGPSGYDNGIITWATWRRRYSKMSV 406

QY 481 SMKIRP 486
 DB 407 TWKIMP 412

RESULT 10
 FGLMGS
 fibrinogen gamma chain precursor - sea lamprey
 C:Species: Petromyzon marinus (sea lamprey)
 C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
 R:Strong, D.D.; Moore, M.; Cottrell, B.A.; Bohonus, V.L.; Pontes, M.; Evans, B.; Riley, A.
 Biochemistry 24, 92-101, 1985
 A:Title: Lamprey fibrinogen gamma chain: cloning, cDNA sequencing, and general characteri
 A:Reference number: A03129; MUID:85199776; PMID:2581603
 A:Accession: A03129
 A:Molecule type: mRNA
 A:Residues: 1-432 <STR>
 A:Cross-references: UNIPROT:P04115; GB:K03049; NID:G213193; PIDN:AAA49262.1; PID:G213194
 C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
 C:Keywords: blood coagulation; glycoprotein; liver; plasma; pyroglutamic acid
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-432/Product: fibrinogen gamma chain #status experimental <MPr>
 F:175-411/Domain: fibrinogen beta/gamma homology <FBG>
 F:25/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
 F:227/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:421/Cross-link: isopeptide (Gln) (interchain to Lys-425 N6-amino) #status predicted
 F:425/Cross-link: isopeptide (Lys) (interchain to Gln-421) #status predicted

Query Match 22.4%; Score 598.5; DB 1; Length 432;
 Best Local Similarity 33.4%; Pred. No. 2.5e-32;
 Matches 144; Conservative 80; Mismatches 138; Indels 69; Gaps 17;

QY 84 KVERKAPDAGGCLHADPDLGVLPTGQLOEALLOQERPIRNSVDELNNVEAVSQTSSS 143
 DB 21 QVHAQVRDLKQCSN-DPEFGRYCPTTCGVADVLSKYAK-----GVEDSSSFISV----- 69

QY 144 SFQYVYLLKDLWKRQKQ--VKDNNVNVNVEVSSELEKHLYI-----DETNSNIATNLRL 196
 DB 70 -----LTQAAKHGIVEGNVIVNE-DVIRTRDEAQIKDSGQKTV-QKILEEVR 117

QY 197 VLRSILENLSKIQKLE-----SDVSAQMEYCRTPCTVSC-----NIPVVSQKE 240
 DB 118 ILQIGVSHDAQISELSEMRVNOQFVTRLOQLVDIQTCSRSQDPTANKISIPITGKD 177

QY 241 CBEIIRKGETSEMYLIQDSSVVKPYRVYCDMNTENG--GWTVIQNRQDGSVDGRKWDVPY 299
 DB 178 CQOVVDNGKSGSLYYIKPLKAKQOFLVYCEI--ENGNGWTVIQHRHDSVNFTRDWSY 235

QY 300 KQGFNVATNTDGRKNYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVAHYGGF 359
 DB 236 REGGYLAPT-----LTTEFWLGNKXIHLLTQQAQVRLRLDLTDWENTHRYADYGHF 287

N;Contains: fibrinopeptide A
C;Species: Homo sapiens (man)
C;Date: 10-Jun-1993 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004
C;Accession: D44234; B44234
R;Pu, Y.; Weissbach, L.; Plant, P.W.; Oddoux, C.; Cao, Y.; Liang, T.J.; Roy, S.N.; Redman
Biochemistry 31, 11968-11972, 1992
A;Title: Carboxy-terminal-extended variant of the human fibrinogen alpha subunit: a novel
A;Reference number: A44234; MUID:93090725; PMID:1457396
A;Accession: D44234
A;Status: translated from GB/EMBL/DDBB
A;Molecule type: mRNA; DNA
A;Residues: 1-866 <F02>
A;Cross-references: UNIPROT:P02671; GB:M58569; NID:g182406; PID:g182407
A;Note: neither the complete nucleic acid sequence nor the complete translation are shown
A;Accession: B44234
A;Molecule type: mRNA; DNA
A;Residues: 605-866 <F02>
A;Note: sequence extracted from NCBI backbone (NCBIP:119917)
C;Comment: The alpha chain binds by 2-4 cross-links to the amino end of fibrinectin.
C;Comment: The conversion of fibrinogen to fibrin is triggered by thrombin, which cleaves
ization sites responsible for the formation of the soft clot.
C;Comment: The soft clot is converted into the hard clot by factor XIIIa (fibrin-stabiliz-
er) and between alpha chains (weaker) of different monomers.
C;Comment: All fibrinogen chains are synthesized in the liver.
C;Comment: See PIR:FGHUA for the major splice form. It is not known whether this form is
C;Genetics:
A;Gene: GDB:FGA
A;Cross-references: GDB:119129; OMIM:134820
A;Map position: 4q28-q28
A;Introns: 18/3; 60/3; 122/1; 171/2
A;Note: The list of introns is incomplete
C;Complex: The fibrinogen molecule is a hexamer containing two sets of three nonidentical
ntained in the core. Two three-chain coiled coils emerge from this core and connect it to
distal domain nodes.
C;Function:
A;Description: fibrinogen cleaved by thrombin yields monomers that are polymerized into
A;Pathway: blood coagulation
C;Superfamily: human extended splice form fibrinogen alpha chain; fibrinogen beta/gamma
C;Keywords: alternative splicing; blood coagulation; glycoprotein; liver; phosphoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-863/Product: fibrinogen alpha chain, extended splice form #status predicted <MAT>
F;20-35/Product: fibrinopeptide A #status experimental <APT>
F;36-863/Product: fibrin alpha chain, extended splice form #status predicted <FGA>
F;57-185/Domain: fibrinogen disulfide ring homology <FDR>
F;591-593/Region: cell attachment (R-G-D) motif
F;629-863/Domain: fibrinogen beta/gamma homology <FBG>
F;22,460/Binding site: phosphate (Ser) (covalent) #status experimental
F;35-36/Cleavage site: Arg-Gly (thrombin) #status experimental
F;47/Disulfide bonds: interchain (to alpha-47) #status experimental
F;55/Disulfide bonds: interchain (to beta-95) #status experimental
F;64/Disulfide bonds: interchain (to gamma-49) #status experimental
F;68/Disulfide bonds: interchain (to beta-106) #status experimental
F;180/Disulfide bonds: interchain (to gamma-165) #status experimental
F;184/Disulfide bonds: interchain (to beta-223) #status experimental
F;288,419/Binding site: carbohydrate (Asn) (covalent) #status absent
F;322/Cross-link: isopeptide (Lys) (interchain to Gln-41 of alpha-2-plasmin inhibitor) #
F;347,385/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of alpha) #status exp
F;461-491/Disulfide bonds: #status experimental
F;527,558,575,591,599/Cross-link: isopeptide (Lys) (interchain to Gln of alpha) #status
F;686,831/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.3%; Score 544.5; DB 2; Length 866;
Best Local Similarity 39.4%; Pred. No. 2,6e-28;
Matches 113; Conservative 58; Mismatches 95; Indels 21; Gaps 8;
QY 207 SKIQKLESDVSQMEYCRTPCTVSCNIPVWSKCEBIIIR--KGSETSEMYLIQDSSVK 264
DB 598 SKSYKVADEAGSEADHEGTHSTKRGHAKSRPVRDCCDLVQTHPSQTSGIFNKLPGSSK 657
QY 265 PVRVYCDMTENGWTVLQNRDGGSDVGRKWDPPYKQCGVATNTDGNKYCGLPGEYWL 324
DB 658 IFSVYCDQETSIGGWLIIQRRDGGSLNFRNTWDYKRGFGSL--NDEGE-----GEFWL 709

QY 325 GNDKISQLTRMGPTTELLIEMEDWKGVKAHYGGFTVQNEANKYQISVKNYRGTFAGNALM 384
DB 710 GNDYLLTLTQRG--SVLRVLELDWAGNEAYAEY-HFRVGSSEAGYALQVSSYEGTAGDALI 767
QY 385 DGASQLMGNRTWTTHNGMFFSYDRDNDGLTSDPRKQCKSDGGGWWYNRCHANNPG 444
DB 768 EGSVE---EGAETSHNNMQFTSFRDADQW-----EENCAEVYGGWYNNCQANLNG 819
QY 445 RYWGQGY-TWDMAKHGTTDDGVVMMNKGWSYMRKSMKIRPFPPQ 490
DB 820 IYYPGGSYDPRNNSPYEINGVWVSFRGADYSLRAVRMKIRLVLVQ 866
RESULT 14
JC8027
type 1 angiotensin II receptor associated protein 1 (ARAP1) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Apr-2004 #sequence_revision 04-Apr-2004 #text_change 04-Apr-2004
C;Accession: JC8027
R;Guo, D.F.; Chenier, I.; Tardif, V.; Orlov, S.N.; Inagami, T.
Biochem. Biophys. Res. Commun. 310, 1254-1265, 2003
A;Title: Type 1 angiotensin II receptor-associated protein ARAP1 binds and recycles the
A;Reference number: JC8027; PMID: 14559250
A;Accession: JC8027
A;Molecule type: mRNA
A;Residues: 1-493 <GUO>
A;Cross-references: GB:AF159049
C;Comment: This protein, which interacts with the carboxyl terminal region of the AT1A re-
nd specifically promotes the recycling of AT1 receptor back to the plasma membrane after
C;Genetics:
A;Gene: arap1
C;Keywords: angiotensin II receptor; receptor recycling
Query Match 19.8%; Score 529.5; DB 2; Length 493;
Best Local Similarity 30.3%; Pred. No. 1,2e-27;
Matches 136; Conservative 69; Mismatches 117; Indels 127; Gaps 15;
QY 112 LOEALIQERPR--NSVDELNNV-----EAVSQTSSSSFOYMYLLKDLWKQK 160
DB 92 LNNELLKQKQETTLQQLVEDVGGIVSEVKLVKESRNNNSRVTLQYMLQLHEIIRKRN 151
QY 161 QVKDNENVVYSSSELEKHQLYIDETVNSNIATNLRLVLSILENLSKIQLK-----ES 214
DB 152 ALK-----LSQLENRL-----NQADMLQVSKYKDLHKPQHLMDLAHNS 194
QY 215 DVSAQM-EYCR-----TPCT-----VSCN----- 232
DB 195 EVIAQLBHCORVPAARVPVPPATPRVYPTVYNNRIINQISTNEIQSDQNLKVLPPS 254
QY 233 IPVV-----SG--KECEIIRKGETSEMILIQDSSVKPYRYVCDMTENG 277
DB 255 IFTPALTSLSSTDKPSGPRDCLQALDGHSTSIYLVKPNTRNLRLMQVWCDQRHDPG 314
QY 278 GMTVIONRQDGSVDPRKWDPPYKQCGVATNTDGNKYCGLPGEYWLGNDKISQLTRMG 337
DB 315 GMTVIOQLDGSVNFRRWYKQGFNI-----DGEVWLGLENLYWITNQGN 362
QY 338 TELIEMEDWKGVKAHYGGFTVQNEANKYQISVKNYRGTFAGNALMGASQLMNGENRTM 397
DB 363 YKLIVTMEDWGRKVFAYASFRLEPESEYKILGRVYHNGAD-----SF 408
QY 398 TIHNGMFFSTYDRDNDGLTSDPRKQCKSDGGGWWYNRCHANNPGRYNGGYQVTDMA 457
DB 409 TWHNGKQFTLLDRHDVYVYTN-----CAHYQKGGWYNNACAFHSLNGLVYRGGHY----- 458
QY 458 KHGTTDDGVVMMNKGWSYMRKSMKIRP 486
DB 459 RSRVDGYWAEFRGGSYSLKXVVMIRP 487
RESULT 15
JN0596
fibrinogen-related protein HFREP-1 precursor - human

RX	MEDLINE=21361164; PubMed=11468164;
RA	Lounes K.C., Lefkowitz J.B., Henschen-Edman A.H., Coates A.I.,
	Query Match 99.8%; Score 2671; DB 1; Length 491;
	Best Local Similarity 99.8%; Pred. No.2.3e-170; Indels 0; Gaps 0;
	Matches 490; Conservative 0; Mismatches 1;
QY	1 MKRMVSNFSFHKLKTMTKMLLLLLLVCFLVKSQVNDNEEGFFSARGHRPLDKKREAFSLR 60
DB	1 MKRMVSNFSFHKLKTMTKMLLLLLLVCFLVKSQVNDNEEGFFSARGHRPLDKKREAFSLR 60
QY	61 PAPPPISGGYRAPPAPAKAAATQKKVERKAPDAGGCIHADPDGLVLCPTGCOLQOALLQOE 120
DB	61 PAPPPISGGYRAPPAPAKAAATQKKVERKAPDAGGCIHADPDGLVLCPTGCOLQOALLQOE 120
QY	121 RPIRNSVDELNNVNAVVSQTSSSSFQWYLLKDLWQKQKQVKNENNVNVEYSELEKQ 180
DB	121 RPIRNSVDELNNVNAVVSQTSSSSFQWYLLKDLWQKQKQVKNENNVNVEYSELEKQ 180
QY	181 LYIDETVNSNIATNLRLVLRSLLENLRSKIQKLESVSAQMEYCRTPCTVSCNIPVVSKE 240
DB	181 LYIDETVNSNIATNLRLVLRSLLENLRSKIQKLESVSAQMEYCRTPCTVSCNIPVVSKE 240
QY	241 CEELIRKGGTSEMYLTQDSSVKPYVYCDMNTENGWTVIQNRQGSVDGFRKWDPPYK 300
DB	241 CEELIRKGGTSEMYLTQDSSVKPYVYCDMNTENGWTVIQNRQGSVDGFRKWDPPYK 300
QY	301 QGFQNVATNTDGKNYCOLPEYMLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFT 360
DB	301 QGFQNVATNTDGKNYCOLPEYMLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFT 360
QY	361 VQNEANKYQISVKNYRGRTAGNALMDGASQWLMGENRTWTIHNGMFFSTYDRNDGWLTS DP 420
DB	361 VQNEANKYQISVKNYRGRTAGNALMDGASQWLMGENRTWTIHNGMFFSTYDRNDGWLTS DP 420
QY	421 RKQCSKEDGGWYNRCHAAHPNGRYTWGQGYTWDMAKHGTDDGVTVMNWKGSYNSRMK 480
DB	421 RKQCSKEDGGWYNRCHAAHPNGRYTWGQGYTWDMAKHGTDDGVTVMNWKGSYNSRMK 480
QY	481 SMKIRPPFPQQ 491
DB	481 SMKIRPPFPQQ 491
RESULT 2	
Q8K0E8	
ID	Q8K0E8 PRELIMINARY; PRT; 481 AA.
AC	Q8K0E8;
DT	01-OCT-2002 (TEMBLrel. 22, Created)
DT	01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT	01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE	Fibrinogen, B beta polypeptide.
GN	Name=Fgb;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RC	SEQUENCE FROM N.A.
RC	STRAN=FVB/N; TISSUE=Liver;
RX	MEDLINE=22388257; PubMed=12477932;
RA	Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Ustin T.B., Tothiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA	Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.
RA	Pahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.,
RA	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences."
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FVB/N; TISSUE=Liver;
RA	Strasbourg R.;
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC031715; AAH31715.1; -
DR	HSP; P02675; 1FEZ.
DR	MGD; MGI:99501; Fgb.
DR	InterPro: IPR002181; Fibrinogen C.
DR	Pfam; PF00147; Fibrinogen_C; 1_
DR	SMART; SM00186; Fbg; 1.
DR	PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ	SEQUENCE 481 AA; 54752 MW; 9902830CF708A155 CRC64;
Query Match 84.4%; Score 2259; DB 2; Length 481;	
Best Local Similarity 83.8%; Pred. No. 8.4e-143;	
Matches 403; Conservative 42; Mismatches 32; Indels 4; Gaps 2	
Qy	15 MKHL-LLLLLLCVFLVKSQGVNDNEE---GPFASGRPLDKKREAPSLRPAPPPISSGG 70
Db	1 MRHLWLLLLLCVSVQTAADDYDEFTDSLDAGRHPVDRRKEPPSLRPAPPPISSGG 60
Qy	71 YRARPAAKAAATQKKVERKAPDAGGCLHADPLGLVCLPTGCOLQEAALQOERPIRNSVDL 130
Db	61 YRARPAKATAQKKVERRPPDAGGCLHADTDMGVLCTGTCTLQTLNQRPIKSSIAEL 120
Qy	131 NNNVEAVSQTSSSSFQMYLLKDLWQKQKQVKNENNVNVIYSSLELKHQLYIDETVNSN 190
Db	121 NNNIQSVSDTSSTVFQYLTLLKDWKKKQKQVKNENVINIYSSILEDQRLYIDETVNDN 180
Qy	191 IATNLRLVRLSILENLRSKIOKLESVSAQMEYCYCTPTCVSCNIPVSGKECEEIIRKGE 250
Db	181 IPLNLRVRLSILEDLRSKIOKLESISQAQMEYCYCTPTCVSCNIPVSGKECEEIIRKGE 240
Qy	251 TSEMYLIQPDSSVKPYRYVYCDMNTENGWTVIQNRQDGSVDFGRKWDPKYKQGFNVATNT 310
Db	241 TSEMYLIQPDTSIKPYRYVCDMTKENGWTVIQNRQDGSVDFGRKWDPKYKGFNVATNE 300
Qy	311 DGKNYCGLPGEYMLGNDKISQLTRMGTELLIEMDWKGVKAHYGGFTVQNEANKYQI 370
Db	301 DAKKYCGLPGEYMLGNDKISQLTRMGTELLIEMDWKGVKAHYGGFTVQNEASKYQV 360
Qy	371 SVNKYRGTAGNALMDGASQLMGENRTWTIHNGMFFSTYDRNDGWLTSDRKQCSKEDGG 430
Db	361 SVNKYKGTAGNALMDGASQLVGENRTWTIHNGMFFSTYDRNDGQWVTTPDRQCSKEDGG 420
Qy	431 GWWYNRCHAAANPGRYYWGQYTWDMAKHGTDGCVVWNNWKGWSYSMRKMSMKIRPFPPQ 490
Db	421 GWWYNRCHAAANPGRYYWGGLYSDWMSKHGTDGCVVWNNWKGWSYSMRMSMKIRPFPPQ 480
Qy	491 Q 491
Db	481 Q 481
RESULT 3	
FIBB_RAT	STANDARD; PRT; 479 AA.
ID_FIBB_RAT	AC P14480;
DT	01-JAN-1990 (Rel. 13, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Fibrinogen beta chain precursor [Contains: Fibrinopeptide B].
OS	Fibrinogen beta chain precursor [Contains: Fibrinopeptide B].
GN	Name=Fgb;
OS	Rattus norvegicus (Rat).


```

DR EMBL; AY321323; AAP86255.1; -.
DR EMBL; AY325153; AAP92554.1; -.
DR EMBL; AY325147; AAP92548.1; -.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; Fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN AG C DOMAIN; 1.
DR SEQUENCE 503 AA; 56612 MW; 02B6A1E05E7B5C12 CRC64;

Query Match 77.7%; Score 2080; DB 2; Length 503;
Best Local Similarity 80.7%; Pred. No. 8.5e-131;
Matches 369; Conservative 42; Mismatches 44; Indels 2; Gaps 1;

QY 34 NDNEGGFSARHRLDCKREAPSLRPAPPISGGYRARPAAKAAATQKKVERKAPDAG 93
DB 21 NSKRDLSTARHRLPVRDRKEPPLRPAPPISGGYRARPAAKADQKKVERKPPDAG 80

QY 94 GCLHADPDLGVLCPGTCQLEALLQOEALQOEPIRNSVDLNNNVEAVSQTSSSFQYMLLKD 153
DB 81 GCVHGDGMVLCPTGCELRLQTLNHERPIKNSIAELNSINSVSETSVTFQYLLTKMD 140

QY 154 LWOKRQKQVKNENNVNVEYSSELEKHQYIDETVNSNATNLRLVLSILENLSKIQLE 213
DB 141 MWKKQQAQYKDNENVINEYSSELEKHQYIDETVNDNPLNRLVLSILENLSKIQLE 200

QY 214 SDVSAQMEYCRTPCTVSCNIPVVSQKCEEEIIRKGETSEMYLIOPDSSVKPYRYCDMN 273
DB 201 SDISAQTEYCHTPCTVNCNIPVVSQKCEEEIIRKGETSEMYLIOPDTSKPYRYCDMK 260

QY 274 TENGWTVIQRQDGSVDFGRKWDPKYQGFNVATNDGKNVCGLPGEYWLGNDKISQIT 333
DB 261 TENGWTVIQRQDGSVDFGRKWDPKYQGFNVATNDGKNVCGLPGEYWLGNDKISQIT 320

QY 334 RMGPTELLIEMEDWKGDKVKAHGGFTVQNEANKYQISVKNYRGTAAGNALMDGASOLMGE 393
DB 321 RIGPTELLIEMEDWKGDKVKAHGGFTVQNEANKYQISVKNYRGTAAGNALMDGASOLVGE 380

QY 394 NRTMTTHNGMFFSTYDRNDGMLTSDPRKQSKEDGGWYRNCHAAANPNRYWGQYT 453
DB 381 NRTMTTHNGMFFSTYDRNDGMLTSDPRKQSKEDGGWYRNCHAAANPNRYWGGLYS 440

QY 454 WDMAKHGTDGDDGVVWNNKGSWYSMRKMSMKIRPFPPQ 490
DB 441 WDMSKHGTDDGVVWNNKGSWTPRQHLQLQRPD--PQ 475

RESULT 6
FIBB_CHICK STANDARD; PRT; 463 AA.
AC Q02020;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Fibrinogen beta chain precursor [Contains: Fibrinopeptide B]
DE (Fragment).
DE Name=FBG;
GN Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-13 AND 18-39.
RP MEDLINE=91182745; PubMed=2009266;
RX Weisbach L., Oddoux C., Procyk R., Grieninger G.;
RA "The beta chain of chicken fibrinogen contains an atypical thrombin
RT cleavage site.";
RL Biochemistry 30:3290-3294 (1991).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: Hexamer containing 2 sets of 3 nonidentical chains
CC (alpha, beta and gamma), linked to each other by disulfide bonds.

```

```

-!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
which cleaves fibrinopeptides A and B from alpha and beta chains,
and thus exposes the N-terminal polymerization sites responsible
for the formation of the soft clot. The soft clot is converted
into the hard clot by factor XIIIa which catalyzes the epsilon-
(gamma-glutamyl)lysine cross-linking between gamma chains
(stronger) and between alpha chains (weaker) of different
monomers.
-!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; M58514; AAA48770.1; -.
PIR; A38463;
PDB; 1BI3; X-ray; B/E=1-463.
PDB; 1MIJ; X-ray; B/E=1-463.
InterPro; IPR002181; Fibrinogen_C.
Pfam; PF00147; Fibrinogen_C; 1.
SMART; SM00186; FBG; 1.
PROSITE; PS00514; FIBRIN AG_C DOMAIN; 1.
3D-structure; Blood coagulation; Direct protein sequencing;
Glycoprotein; Plasma; Platelet; Sulfation.
NON_TER 1
PEPTIDE <1 17 Fibrinopeptide B.
CHAIN 18 463 Fibrinogen beta chain.
MOD_RES 5 5 Sulfotyrosine (By similarity).
SITE 17 18 Cleavage (by thrombin; release
fibrinopeptide B).
DISULFID 69 69 Interchain (with alpha) (By similarity).
DISULFID 80 80 Interchain (with alpha) (By similarity).
DISULFID 84 84 Interchain (with gamma) (By similarity).
DISULFID 197 197 Interchain (with alpha) (By similarity).
DISULFID 201 201 Interchain (with gamma) (By similarity).
DISULFID 205 289 By similarity.
DISULFID 215 244 By similarity.
DISULFID 397 410 By similarity.
CARBOHYD 367 367 N-linked (GlcNAc...) (Potential).
SEQUENCE 463 AA; 52678 MW; 2044CD49BA79EC7B CRC64;

Query Match 65.2%; Score 1744; DB 1; Length 463;
Best Local Similarity 66.2%; Pred. No. 2.3e-108;
Matches 305; Conservative 70; Mismatches 76; Indels 10; Gaps 4;

QY 35 DNEEGF--FSARGHRLDCKREAPSLRPAPPISGGYRARPAAKAAATQKKVERKA--- 89
DB 6 DNEEDSFQIDARAHRLDCKREAPSLRPAPPISGGYRARPAAKAAATQKKVERKA--- 62
QY 90 -PDAGGLHADPDLGVLCPGTCQLEALLQOEALQOEPIRNSVDLNNNVEAVSQTSSSFQYM 148
DB 63 YPDAGGCKHPLDELGLVLCPTGCELOTLKQEKTVKPELDELKDRVAKFSDTSTTWYQYV 122

QY 149 YLLKDLWQRQKQVKNENNVNVEYSSELEKHQYIDETVNSNATNLRLVLSILENLSK 208
DB 123 NMIDNKLVTQKQKNDIILSEYNTSEMELHNYIKDNLNDNTPSSIRVLRVRAVDSLHK 182

QY 209 IQKLESDVAOMEYCTPTCTVSCNIPVVSQKCEEEIIRKGETSEMYLIOPDSSVKPYRY 268
DB 183 IQKLENAIATQTDYCRSPCVASCNIPVVSQKCEEEIIRKGETSEMYLIOPDFTFYRV 242

QY 269 YCDMTENGWTVIQRQDGSVDFGRKWDPKYQGFNVATNDGKNVCGLPGEYWLGNDK 328
DB 243 YCDMTENGWTVIQRQDGSVDFGRKWDPKYQGFNVATNDGKNVCGLPGEYWLGNDK 301

QY 329 ISQITRMGPTELLIEMEDWKGDKVKAHGGFTVQNEANKYQISVKNYRGTAAGNALMDGAS 388
DB 302 ISQITKIGTPKVLIEEMEDWNGDKVSAFYGGFTIHNEGKQYLSVSNYKGNAGNALMEGAS 361

```


QY	389	OLMGENTMTIHNGMFSTYDRDNDGWLTSDBRKQCSKEDGGGWNYNRCHAAANPNGRYYW	444
Db	362	QLYGENTMTIHNGMYFSTYDRDNDGWLTTDPKQCSKEDGGGWNYNRCHAAANPNGRYYW	421
QY	449	GGQYTDWAKHGTDGIVVMWNGKSWYSMRKSMKIRPFPP	489
Db	422	GGYSDWAKHGTDGIVVMWNGKSWYSMRKSMKIRPFPP	462
RESULT 7			
Q91589 PRELIMINARY; PRT; 488 AA.			
ID	Q91589	Q91589; Q91546;	
AC	Q91589; Q91546;		
DT	01-NOV-1996 (TrEMBLrel. 01, Created)		
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Fibrinogen B-beta subunit precursor.		
DE	Fibrinogen B-beta subunit precursor.		
OS	Xenopus laevis (African clawed frog).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;		
OC	Xenopodinae; Xenopus.		
OC	NCBI_TaxID=8355;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RX	MEDLINE=95369693; PubMed=7642099;		
RA	Roberts L.R., Nichols L.A., Holland L.J.;		
RT	"cDNA and amino-acid sequences and organization of the gene encoding		
RT	the B beta subunit of fibrinogen from Xenopus laevis.";		
RL	Gene 160:223-228(1995).		
RL	EMBL; U19618; AAA85283.1; -.		
DR	PIR; A48810; A48810.		
DR	HSP; P02675; ILTJ.		
DR	InterPro; IPR002181; Fibrinogen_C.		
DR	Pfam; PF00147; Fibrinogen_C; 1.		
DR	SMART; SM00186; FBG; 1.		
DR	PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.		
KW	Signal.		
FT	CHAIN	1 16	
FT	CHAIN	17 43	B fibrinopeptide.
FT	CHAIN	17 488	fibrinogen B-beta subunit.
FT	SEQUENCE	488 AA; 54803 MW; 660E03844BBC6414 CRC64;	
Query Match		64.8%; Score 1734.5; DB 2; Length 488;	
Best Local Similarity		63.6%; Pred. No. 1.1e-107;	
Matches 313; Conservative		92; Mismatches 92; Indels 23; Gaps 4;	
QY	15	MKHLILLLCV-----FLVKSQGVNDNEGFSARGHRPLDKKREAP	57
Db	1	MRVLLLFALCVAVWCSSDYDEDDVDAAVIKS----DNATASVDARGHRPVSGRFPV	56
QY	58	SLRPAPPPISSGGYRARPAKAAAT-QKKVERKAPDAGGLHADPDLVLCPTGCGLOEAL	116
Db	57	TQRPAPPPISSGGYRGPRPTKAPVGOKKEATEYDAGCKHAYEELGTLCPTGCELTTTL	116
QY	117	LQERPIRNSVDLNNVAVSQTSSSFQWYLLKDLWKQKQKQKDNENNVNVEYSEL	176
Db	117	LKQERNVKTAINDDVRGVRVETLAQSANNVRYTTLVGQKI KENQQTLDNQNVVNEYNLEL	176
QY	177	EXHQLYIDEVNSIATNLRLVLSILENLSKQKLESVDVSAQMEYCRTPCTVSCNIPV	236
Db	177	EQYTFIKONIDTKPSNIRILRQVLENLSKIQKLETAATQVENCSPCVTTTCPIPV	236
QY	237	SGKECEBIIRKGGTSEMYLIQPDSSVKKPYRVYCDMNTENGWTVIQRNODGSDVDFGRKW	296
Db	237	SGKECEBIIRKGGTSEMYLIQPDSSFRPPKVKYCDMATHDGGWTVIQRNODGSGVFGRTW	296
QY	297	DPYKQGFNVATNTDGKNYCGLPGEYWLGNDKISQLTRMGPTLELLIEMEDWKDGKVAHY	356
Db	297	DSYKSGFGNTAAN-GGKGI CDMPGEFWLGNKISQLTNLGATEALFEMEDWDGAKVTAQY	355
QY	357	GGTIVQNEANKYQISVNYKYGRTAGNALMDGASQLMGENRTWTIHNGMFFSTYDRDNDGWL	416

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2].
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3].
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA Klein S., Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC054326; AAH54326.1; -;
 DR InterPro; IPR002181; Fibrinogen_C.
 DR Pfam; PF00147; Fibrinogen_C; 1.
 DR SMART; SM00186; FBG; 1.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
 FT NON TER 1
 SQ SEQUENCE 321 AA; 36519 MW; E0C41FCC3EB40DEF CRC64;

 Query Match 51.7%; Score 1384.5; DB 2; Length 321;
 Best Local Similarity 75.4%; Pred. No. 1.6e-84;
 Matches 242; Conservative 33; Mismatches 45; Indels 1; Gaps 1;

 QY 168 VVNEYSSELEKQLYIDETVNSNIATNLRLVLSILENLRSKIQLSDVSQAQMEYCRTPC 227
 DB 1 VVNEYNLEEEQYTFIKDNIIDTKIPSNIRILRQVLENLSRKIKLTAIATQVENCSPC 60

 QY 228 TVSCNIPVVGKCECEIIRKGETSEMYLIQDSSVVKPVYVCDMTENGWTVIQRQD 287
 DB 61 VTTCPIPVVGKCECEIIRKGETSEMYLIQDSSVVKPVYVCDMTENGWTVIQRQD 120

 QY 288 GSVDFGRKWDPKYKQGFNGVATNTDGNKYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDW 347
 DB 121 GSVGFGRTWDSYKSGFGNTAAN-GKGICDMGFEFWLGNKEISQLINLGATLAFEMEDW 179

 QY 348 KGDVKVAHYGGFTVQNEANKYQISVKNYKGTAGNALMDGASQLMGENRTMTIHNGMFFST 407
 DB 180 DGAKVTAQYTGFTVQNEANKYQISVKNYKGTAGNALMDGASQLMGENRTMTIHNGMFFST 239

 QY 408 YDRDNDGMLTSDPRKQCKEDGGWYNNRCHAANPNRGYVGGQYTWDMAKHGTDDGVVW 467
 DB 240 FDRDNDGWHQSDPNKQCKEDGGWYNNRCHAANPNRGYVGGQYTWDMAKHGTDDGVVW 299

 QY 468 MNWKGWSYMSRMKIRPFF 488
 DB 300 MNWKGWSYMSRMKIRPFF 320

 RESULT 11
 Q91ZP1 PRELIMINARY; PRT; 236 AA.
 AC AC
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-WAR-2004 (TREMELrel. 26, Last annotation update)
 DE Fibrinogen B-beta-chain (Fragment).
 GN Name=FBG;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;

RN SEQUENCE FROM N.A.
 RP Murakawa M., Freeman M.W.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF413205; AAL02225.1; -;
 DR MGD; MGI:99501; Fgb.
 DR InterPro; IPR002181; Fibrinogen_C.
 DR Pfam; PF00147; Fibrinogen_C; 1.
 DR SMART; SM00186; FBG; 1.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
 FT NON TER 1
 SQ SEQUENCE 236 AA; 27054 MW; A844F0E40CCB8B01 CRC64;

 Query Match 47.8%; Score 1279; DB 2; Length 236;
 Best Local Similarity 92.4%; Pred. No. 1.2e-77;
 Matches 218; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

 QY 244 IIRKGETSEMYLIQDSSVVKPVYVCDMTENGWTVIQRQDGSVDFGRKWDPKYKQGF 303
 DB 1 IIRKGETSEMYLIQDSSVVKPVYVCDMTENGWTVIQRQDGSVDFGRKWDPKYKQGF 60

 QY 304 GNVAITNTDGNKYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQN 363
 DB 61 GNIATNEDAKYTCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQN 120

 QY 364 EANKVQISVKNYKGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGMLTSDPRKQ 423
 DB 121 EASKYQVSVNKKYKGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGMLTSDPRKQ 180

 QY 424 CSKEDGGGWWYNNRCHAANPNRGYVGGQYTWDMAKHGTDDGVVWNNWKGWSYMSRMK 479
 DB 181 CSKEDGGGWWYNNRCHAANPNRGYVGGQYTWDMAKHGTDDGVVWNNWKGWSYMSRMK 236

 RESULT 12
 FIBB PETWA
 ID FIBB PETWA STANDARD; PRT; 477 AA.
 AC P02678;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragments).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OC NCBI_TaxID=7757;
 RN [1].
 RP SEQUENCE OF 1-36.
 RX MEDLINE=7065679; PubMed=998988;
 RA Cottrell B.A., Doolittle R.F.;
 RT "Amino acid sequences of lamprey fibrinopeptides A and B and
 RT characterizations of the junctions split by lamprey and mammalian
 RT thrombins.";
 RL Biochim. Biophys. Acta 453:426-438(1976).
 RN [2].
 RP SEQUENCE OF 37-477 FROM N.A.
 RX MEDLINE=87076582; PubMed=3790537;
 RA Bohonus V.L., Doolittle R.F., Pontes M., Strong D.D.;
 RT "Complementary DNA sequence of lamprey fibrinogen beta chain.";
 RL Biochemistry 25:6512-6516(1986).
 CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
 CC polymerize into fibrin and acting as a cofactor in platelet
 CC aggregation.
 CC -!- SUBUNIT: Hexamer containing 2 sets of 3 nonidentical chains
 CC (alpha, beta and gamma), linked to each other by disulfide bonds.
 CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
 CC which cleaves fibrinopeptides A and B from alpha and beta chains,
 CC and thus exposes the N-terminal polymerization sites responsible
 CC for the formation of the soft clot. The soft clot is converted
 CC into the hard clot by factor XIIIa which catalyzes the epsilon-
 CC (gamma-glutamyl)lysine cross-linking between gamma chains
 CC (stronger) and between alpha chains (weaker) of different

Cyprinidae; Danio.
NCBI_taxid=7955;
[1]
SEQUENCE FROM N.A.
STRAIN=AB; TISSUE=whole body;
MEDLINE=22388257; PubMed=12477932;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buotow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.B.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., McQuellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., Leowanawan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzywinski M.I., Skalska U., Smalls D.E., Schnercher A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
SEQUENCE FROM N.A.
STRAIN=AB; TISSUE=Whole body;
Strausberg R.;
Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
EMBL; BC045868; AAH45868.1; -.
HSP; P02679; 2FIB.
InterPro; IPR002181; Fibrinogen_C.
Pfam; PF00147; Fibrinogen_C_1.
SMART; SM00186; FBG_1
PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SEQUENCE 431 AA; 48832 MW; 101B2E47E33DFF71 CRC64;
Query Match 24.6%; Score 657; DB 2; Length 431;
Best Local Similarity 35.2%; Pred. No. 1.2e-35;
Matches 141; Conservative 85; Mismatches 127; Indels 48; Gaps 13;
QY 103 GVLCPTGCQLQEALLQOERPIRNSVDLNNNVAVSQTSSSFQMYLLKDLWQKRQKQ- 161
Db 39 GTYCPTTCGVADYLQRYKPDMDKKLDMEQDLBEIANLTRGAQDKVYVKDSEAQAQKS 98
QY 162 ----VKDNENVVN----EYSSELEXHOLYIDETVNSNIATNLVLASILENLSKIQL 212
Db 99 PDTYIKKSNMLDDILRFEXSLIAQSEQIY-----QLQSVLQANEKITDL 144
QY 213 ESDVSAQM-EYCRTPCTVSCNIPWSKECEEIIRKGGETSEMVLIOPDSSVPYRVCD 271
Db 145 K-QMSQLDMQCKPECKDTVEIQTVTGKQODIANKGKVGSLGYYPAPAPAEFLVYCE 203
QY 272 MTNTENGWTVIQNRQGSVDFGRKWDPYKOGPGNVATNTDGKNYCGLPGEYWLGNDKISQ 331
Db 204 IDSFGRGVTQLRRRDGSGVDFSKNWIQYKEGFGL--SPDRT-----EFWLGNEKIHL 255
QY 332 LTRWG--PTELLIEMDWKGDVKYKAHVGGFTVQNANKYQISVNYK-RGTAGNALMDGAS 388
Db 256 LSVOSSVPVPLRLMVDWEGNKKYADYATFKLGPFDVAVLTYYAYFGGDAGDA-FDGYD 314
QY 389 QLMGE---NRTMTIHNGMFFSTYDRNDGWLTSDPRKQCSKDEGGWYNRCHAAPNGR 445
Db 315 --FGDDPSDKFYTHSHGMQESTVDRNDKF-----QGHCAQQDGSQWMNRCHAALNGK 367
QY 446 YTWGGYTWDMAKHGTDGCVVMWNKGSWYSMRKSMKIRP 486
Db 368 YYQGQGYTKDAESGYDNGIIIMATWHSRWYSLKETTKIIP 408

Query Match	23.6%	Score	631.5	DB 2	Length	436			
Best local Similarity	33.5%	Pred. No.	6e-34			15			
Matches	140	Conservative	79	Mismatches	138	Indels	61	Gaps	15
Qy	95	CLHADPDLGVLCPTGCQLQEQALLQOQRPINRSVDELNNV-EAVSQTSSS-----SFOYM	148						
Db	33	CCILDERFGSPCTTCGADFLSSYQTDVNDLTKLEDILFRAENRTTEAKELIKATQVY	92						
Qy	149	Y-----LLKDLWQKRQKQVDNENNVNVEYSSELEKHQL-----YIDETVNSIATNL	196						
Db	93	YNPDQPKPGMIDSATQSKWVEE-----IVKYEALLLTHETSIRYLQEIYNSN-----	142						
Qy	197	VILRSILENLRSKIQKLESVDSAQMEYCRTPCTVSCNIPVVGSGKECEETIRKGETSEMYL	256						
Db	143	--NQKITNLKQKVALEAQ-----CQEPCKDSVQIHDTTKDCKQEIANKGAKESGIYF	193						
Qy	257	IQPDSSVKPYRYVCDMNTENGCTWTIQRQDGSVDYFGRKWDPPYKQGFGNVA-TMTDGNKY	315						
Db	194	IRPLKAKQFLVYCEIDGSGNGVTWLQKRIDGSLDFKKNWIKYQKEGFGHLSPGTGTT----	249						
Qy	316	CGLPCEYLNWKNKISQLTFMG--PTELLIEMEDMKGDKVKAHYGGFTVONANKYQISVN	373						
Db	250	-----EFLWLGNEKHILISMQSTIPVALRIQLKQDWMNGRTSDIAYMFRVGPSSDKRLTYA	304						
Qy	374	KY-RGTAGNALMDGASQLMGE---NRTMTIHNGMEFFSTYDRDNDGWL*SDRPKQCKSEDG	429						
Db	305	YFIGGDAGDA-FDGYD--FGDDPSDKFFTSHNGMQFSTWINDNDKF-----EGNCAEQDG	356						
Qy	430	GGWMTYNRCHAANPNGRYYWGGQYTWDMAKGHDTDGVVWNNWKGSIYSMKKMSMKTRPF	487						

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:54:56 ; Search time 14396.5 Seconds
(without alignments)
12349.531 Million cell updates/sec

Title: US-10-017-724-5_COPY_4000_8878

Perfect score: 4879

Sequence: 1 tcagaataatgtgtgatgag.....gattgcattgagccaagatc 4879

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	696.6	14.3	716	5	BX118749
2	686	14.1	768	4	BG563877
3	671.6	13.8	820	7	CO775859
4	662.4	13.6	758	4	BG568684
C 5	655.2	13.4	679	4	BM683513
6	650.4	13.3	673	8	AQ489171
7	644.4	13.2	1891	3	CR594769
8	643.4	13.2	1577	3	CR596720
9	642.4	13.2	1883	3	CR596576
C 10	641.4	13.1	1888	3	CR611717
11	640.4	13.1	1014	5	BX448857
C 12	639.4	13.1	1886	3	CR596719
C 13	637.8	13.1	997	5	BX427240
14	634.4	13.0	1880	3	CR626133
C 15	634	13.0	804	1	AW003933
C 16	634	13.0	1016	5	BX427173
C 17	632.4	13.0	1882	3	CR593920
C 18	628.4	12.9	1038	5	BX448675
19	628.4	12.9	1577	3	CR615740
20	628.4	12.9	1875	3	CR614263
21	627.4	12.9	1873	3	CR609458
C 22	626.2	12.8	1038	5	BX405660
23	625.4	12.8	1872	3	CR604009
24	624.4	12.8	1869	3	CR617322

C 25	624	12.8	937	5	BX346072	BX346072
C 26	624	12.8	1041	5	BX346089	BX346089
C 27	623.4	12.8	1572	3	CR598319	CR598319
28	623.4	12.8	1848	3	CR597172	full-length
C 29	622.6	12.8	1064	5	BX427236	BX427236
C 30	621.4	12.7	1862	3	CR594221	full-length
C 31	621.2	12.7	1045	5	BX448329	BX448329
C 32	618	12.7	930	1	AL564576	AL564576
33	616.4	12.6	1859	3	CR602828	full-length
34	615.6	12.6	699	1	AV652183	AV652183
C 35	615.4	12.6	636	2	AW771537	hm57h10.x
C 36	615	12.6	1007	1	AL579878	AL579878
C 37	614.2	12.6	936	5	BX423808	BX423808
C 38	611.4	12.5	1029	1	AL579779	AL579779
C 39	609.8	12.5	1008	5	BX448288	BX448288
40	608.8	12.5	664	1	AV681947	AV681947
C 41	608.2	12.5	1027	5	BX405881	BX405881
42	607.6	12.5	628	5	BX494484	BX494484
C 43	606.2	12.4	1029	5	BX427659	BX427659
C 44	604.8	12.4	761	1	AV700652	AV700652
C 45	604.6	12.4	1009	1	AL531649	AL531649

ALIGNMENTS

RESULT 1
BX118749/c
LOCUS BX118749 716 bp mRNA linear EST 10-FEB-2003
DEFINITION IMAGp998P03372 ; IMAGE:195698, mRNA sequence.
ACCESSION BX118749
VERSION BX118749.1 GI:27841363
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 716)
AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M., Radloff, U., Schneider, D. and Korn, B.
TITLE Human UnigeneSet - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998P03372.
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.

FEATURES

Location/Qualifiers
1..716
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGp998P03372 ; IMAGE:195698"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer (5' AACTGGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3')."

double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

```
Query Match      14.3%; Score 696.6; DB 5; Length 716;
Best Local Similarity 99.3%; Pred. No. 3.2e-110;
Matches 710; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 3497 ACTAGAATTGAGCACAGTACTTCTGGAATPAATGAGCAATTAGATAAATACCTATATCT 3556
Db |||||||
Qy 716 ACTAGAATTGAGCACAGTACTTCTGGAATPAATGAGCAATTAGATAAATACCTATATCT 657
Db |||||||

Qy 3557 CTATATTTCTAAAGTGTGTGGAACACTTTCATTTCTTCCAGGGTCTCTGATACCTAA 3616
Db |||||||
Qy 656 CTATATTTCTAAAGTGTGTGGAACACTTTCATTTCTTCCAGGGTCTCTGATACCTAA 597
Db |||||||

Qy 3617 GGGTTGTAAGAGCTATTATTCAGTATATAAGTAACAAACACAGTCCCTAGATGATT-GC 3675
Db |||||||
Qy 596 GGGTTGTAAGAGCTATTATTCAGTATATAAGTAACAAACACAGTCCCTAGATGATTGGC 537
Db |||||||

Qy 3676 CACAAGGCCCGAGTTATCTCTTTCTTGTATAGGGCACAGGAGTCTTTGGTGTATTA 3735
Db |||||||
Qy 536 CACAAGGCCCGAGTTATCTCTTTCTTGTATAGGGCACAGGAGTCTTTGGTGTATTA 477
Db |||||||

Qy 3736 GTGTGACTCTATGTATAGCACCAAGGAAGACTTCTGTGCACACAGAGTGTAGCAGTCT 3795
Db |||||||
Qy 476 GTGTGACTCTATGTATAGCACCAAGGAAGACTTCTGTGCACACAGAGTGTAGCAGTCT 417
Db |||||||

Qy 3796 TTTATGGGTAACTCTCAAAAGCTAACTTGACACCGTAGTCTGTCTTCTTAATAAGCCAA 3855
Db |||||||
Qy 416 TTTATGGGTAACTCTCAAAAGCTAACTTGACACCGTAGTCTGTCTTCTTAATAAGCCAA 357
Db |||||||

Qy 3856 ACACATTTCTTTCAGTTAATCATCAGATGCCAGAAAACAGTGTCTTAAAGAACGGTG 3915
Db |||||||
Qy 356 ACACATTTCTTTCAGTTAATCATCAGATGCCAGAAAACAGTGTCTTAAAGAACGGTG 297
Db |||||||

Qy 3916 GTGGATGTGGTATATAGATCTCATGAGCCAAATCCAAACGGCAGATCTACTGGGGTG 3975
Db |||||||
Qy 296 GTGGATGTGGTATATAGATCTCATGAGCCAAATCCAAACGGCAGATCTACTGGGGTG 237
Db |||||||

Qy 3976 GACAGTACACCTGGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGATGAATT 4035
Db |||||||
Qy 236 GACAGTACACCTGGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGATGAATT 177
Db |||||||

Qy 4036 GGAAGGGGTCTATGTTACTCAATGAGGAAGATGAGTATGAAGATCAGGCCCTTCTTCCAC 4095
Db |||||||
Qy 176 GGAAGGGGTCTATGTTACTCAATGAGGAAGATGAGTATGAAGATCAGGCCCTTCTTCCAC 117
Db |||||||

Qy 4096 AGCAATAGTCCCAATAGTATGATTTTCTCTCTGTATGTGACAAACATTTTGTACAT 4155
Db |||||||
Qy 116 AGCAATAGTCCCAATAGTATGATTTTCTCTCTGTATGTGACAAACATTTTGTACAT 57
Db |||||||

Qy 4156 TATGTTATTTGGAATTTCTTTTCATACATTATATTTCTCTTAAACCTCTCAAGCAGA 4210
Db |||||||
Qy 56 TATGTTATTTGGAATTTCTTTTCATACATTATATTTCTCTTAAACCTCTCAAAAAA 2
Db |||||||
```

RESULT 2

BG563877 768 bp mRNA linear EST 10-APR-2001
LOCUS 602584686F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4712588 5',
DEFINITION mRNA sequence.

ACCESSION BG563877

VERSION BG563877.1 GI:13571529

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 768)

AUTHORS

TITLE NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: csaphe-r@mail.nih.gov

Tissue Procurement: CLONETECH Laboratories, Inc.

cDNA Library Preparation: CLONETECH Laboratories, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHCMI554 row: k column: 21

High quality sequence stop: 764.

FEATURES

source

1..768

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="IMAGE:4712588"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC_76"

/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site 1:

SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence:

5'-ATTCTAGAGCCGAGCGGCGCATG-3' (30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.85

kb (range 1.0-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA). Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match      14.1%; Score 686; DB 4; Length 768;
Best Local Similarity 97.8%; Pred. No. 2.2e-108;
Matches 749; Conservative 0; Mismatches 10; Indels 7; Gaps 5;

Qy 554 GAAACAGCTAGTGTGAGCGCTGGCACAGGCAATCTCAGTAGTAGTACATCCACAGTTGG 613
Db |||||||
Qy 1 GAAACAGCTAGTGTGAGCGCTGGCACAGGCAATCTCAGTAGTAGTACATCCACAGTTGG 60
Db |||||||

Qy 614 AAATTTTCAAGAAATCAAGGAATCATGACATCTTATAATTTCAAGGTTCTCTATA 673
Db |||||||
Qy 61 AAATTTTCAAGAAATCAAGGAATCAAGGAATCATGACATCTTATAATTTCAAGGTTCTCTATA 120
Db |||||||

Qy 674 CTTATGTCAAATGGATAAATAAATCAAGCATATCCACTCTGTAAAGATTGAATCTTCAGA 733
Db |||||||
Qy 121 CTTATGTCAAATGGATAAATAAATCAAGCATATCCACTCTGTAAAGATTGAATCTTCAGA 180
Db |||||||

Qy 734 TGGAAAGACCCCAATCTGCTTTCTCTCTTTTCCCTCCAAAGAAATAAACACCTATT 793
Db |||||||
Qy 181 TGGAAAGACCCCAATCTGCTTTCTCTCTTTTCCCTCCAAAGAAATAAACACCTATT 240
Db |||||||

Qy 794 TCATTTATTTCTGGACACAATCTTTAGCGTATACCTATGTTAAATTTACTAGTATGGTGG 852
Db |||||||
Qy 241 TCATTTATTTCTGGACACAATCTTTAGCGTATACCTATGTTAAATTTACTAGTATGGTGG 300
Db |||||||

Qy 853 TTAGGATTTATGTTAAATTTGTATATGTCTATGCGCCAAATCAATTTCCACTTAAATGACTA 912
Db |||||||
Qy 301 TTAGGATTTATGTTAAATTTGTATATGTCTATGCGCCAAATCAATTTCCACTTAAATGACTA 360
Db |||||||

Qy 913 TATATCATAACTGCTGGTGTAGTCTAGTGTGTTTAAATAGTTTATTTCTCGAAAAATCAAAA 972
Db |||||||
Qy 361 TATATCATAACTGCTGGTGTAGTCTAGTGTGTTTAAATAGTTTATTTCTCGAAAAATCAAAA 420
Db |||||||

Qy 973 TTGTATAGTTAAATACATTTAGTTTATCGAGGCAAAATGCTAACTATTTCTACATAATTT 1032
Db |||||||
Qy 421 TTGTATAGTTAAATACATTTAGTTTATCGAGGCAAAATGCTAACTATTTCTACATAATTT 480
Db |||||||

Qy 1033 CAATTTTCCAGATAATGAA---AATGTAGTCAATGAGTACTCTCTCAGAACTGGAAGAAGCA 1089
Db |||||||
Qy 481 CAATTTTCCAGATAATGAAACACAGTGTAGTCAATGAGTACTCTCTCAGAACTGGAAGAAGCA 540
Db |||||||
```


QY 1090 CCAATTATATA-TAGATGAGAC-TGTGAATAGCA-ATATCCCACTAACTTCGTGTGCT 1146
 Db |||||
 541 CCAATTATATA-TAGATGAGAC-TGTGAATAGCA-ATATCCCACTAACTTCGTGTGCT 600
 QY |||||
 1147 TCGTTCATCTCGAAGAACCTGAGAGCAAAATACAAAAGTTAGAATCTGATCTCAGC 1206
 Db |||||
 601 TCGTTCATCTCGAAGAACCTGAGAGCAAAATACAAAAGTTAGAATCTGATCTCAGC 660
 QY |||||
 1207 TCAAAATGGAATATGTCGACCCCATGCTGTCAGTTCGATGCAATATTCCTGTGCTGCTCG 1266
 Db |||||
 661 TCAAAATGGAATATGTCGACCCCATGCTGTCAGTTCGATGCAATATTCCTGTGCTGCTCG 720
 QY |||||
 1267 CAAAGTAACTGATTCATAAACATATTTTATAGAGAGTTCCAGAAGA 1312
 Db |||||
 721 CAAAGTAACTGATTCATAAACATATTTTATAGAGAGTTCCAGAAGA 766

RESULT 3
 CO775659 820 bp mRNA linear EST 04-AUG-2004
 LOCUS ILLUMIGEN MCQ 53006 Katze_MFLV Macaca fascicularis cDNA clone
 DEFINITION IBIUW:28390 57 similar to Bases 199 to 720 highly similar to human
 FBG (H8.300774), mRNA sequence.
 ACCESSION CO775659 1 GI:50976922
 VERSION EST.
 KEYWORDS Macaca fascicularis (crab-eating macaque)
 SOURCE Macaca fascicularis
 ORGANISM Macaca fascicularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Macaca.
 REFERENCE 1 (bases 1 to 820)
 AUTHORS Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magnus, C.L.
 TITLE Large-scale Rhesus Macaque cDNA Sequencing
 JOURNAL Unpublished (2003)
 COMMENT Contact: C. Magnus
 Illumigen Biosciences Inc.
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400
 Fax: 2063780408
 Email: cmagnus@illumigen.com
 Sequenced on 2004.07.29. 657 Q20 bases. Library Preparation: Prof.
 Michael Katze Lab at University of Washington DNA Sequencing:
 Illumigen Biosciences Inc. For further information, see
 http://www.macaque.org
 PCR Primers
 FORWARD: CCTCACTAAAGGGACAA
 BACKWARD: CACTATAGGGGAATTGGTA
 Insert Length: 820 Std Error: 0.00
 Plate: CL000536 row: F column: 07
 Seq primer: CCTCACTAAAGGGACAA
 POLYA=Yes.

FEATURES
 Location/Qualifiers
 1..820
 /organism="Macaca fascicularis"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9541"
 /clone="IBIUW:28390"
 /dev_stage="adult"
 /lab_host="Electromax DH10B"
 /clone_lib="Katze_MFLV"
 /note="Organ: liver; Vector: pDONR 222; Site 1: BsrG I;
 Site 2: BsrG I; Created from Cloneminer cDNA Library
 Construction kit (catalog #18249-029)"

ORIGIN
 Query Match 13.8%; Score 671.6; DB 7; Length 820;
 Best Local Similarity 91.6%; Pred. No. 6.6e-106;
 Matches 744; Conservative 0; Mismatches 64; Indels 4; Gaps 3;
 QY 3672 TTGCCAAAAGCCCGAGTATCTCTCTTTCTGTATAGGGCAGAGGAGTCTTTGGTGT 3731

Db |||||
 1 TTGGCAAAAGGCCAGTCTCTCTCTTTCTTATAGGCGCAGAGGTCTTTGGTAT 60
 QY |||||
 3732 ATTAGTGTGACTCTATGTATAGCACCCAAAGGAAGAGACTCTGTGCACAGAGTAGCA 3791
 Db |||||
 61 ATTAGTGTGACTCTATGTGTAGCACCCAAAGGAAGAGTACTGTGCACAGAGTAGCA 120
 QY |||||
 3792 GTCTTTTATGGGTAACTCGAAAAAGTAACTTGACCCCGTAGTCTCTGTTCTAAATAAG 3851
 Db |||||
 121 GTCTTTTATGGGTAACTCGAAAAAGTAACTTGACCCCGTAGTCTCTGTTCTAAATAAG 179
 QY |||||
 3852 CCAAAACACATTTTCTTTTCAGGTTAACTCAGATCCAGAAAAACAGGTCTTAAAGAGAC 3911
 Db |||||
 180 CCAAAACACATTTTCTTTTCAGGTTAACTCAGATCCAGAAAAACAGGTCTTAAAGAGAT 239
 QY |||||
 3912 GGTGGTGGATGGTGGTATTAATAGATGTATGACAGCAATCCAAACGGCAGATACTACTGG 3971
 Db |||||
 240 GGTGGTGGATGGTGGTATTAATAGTGTACGCGAGCAATCCAAATGCGAGATACTACTGG 299
 QY |||||
 3972 GGTGGACAGTACACCTGGGACATGGCAAGCATGGCAGATGATGGTGTAGTATGGATG 4031
 Db |||||
 300 GGTGGAAAAATATACCTGGGACATGGCAAGCATGGCAGATGATGGGTGTAGTATGGATG 359
 QY |||||
 4032 AATTGGAAGGGTCACTGGTACTCAATGAGAAAGATGAGTATGAAGATCAGGCCCTTCTTC 4091
 Db |||||
 360 AATTGGAAGGGTCACTGGTACTCAATGAGAAAGATGAGTATGAAGATCAGGCCCTTCTTC 419
 QY |||||
 4092 CCACAGCAATAGTCCCAATACGTAGATTTTCTCTCTCTGTATGTGACAAACATTTTGT 4151
 Db |||||
 420 CCACAGCAATAGTCCCAATACGTAGATTTTCTCTCTCTGTATGTGACAAACATTTTGT 479
 QY |||||
 4152 ACATATGTATTGGAATTTCTTTCATACATATATTTCTCTCTCTCTCTCTCTCTCTCTCT 4211
 Db |||||
 480 ACATATGTATTGGAATTTCTTTCATACATATATTTCTCTCTCTCTCTCTCTCTCTCTCT 539
 QY |||||
 4212 GTGAGTGTGACTTTTGAAGAAAGTATAGTAATTAATTAATTAATTAATTAATTAATTAAT 4270
 Db |||||
 540 GTGAGTGTGACTTTTGAAGAAAGTATAGTAATTAATTAATTAATTAATTAATTAATTAAT 599
 QY |||||
 4271 TCTTTTGTCTTCT 4330
 Db |||||
 600 TCTTTTGTCTTCT 659
 QY |||||
 4331 GAGTGTGTTTCATATTTTCAGTTCTAGTGTGAGAGATTTTCAATAAGAGAGAGG 4390
 Db |||||
 660 GAGTGTGTTTCATATTTTCCATTTCTGTTGCTGTAAGTTTCAATAAGAGAGAGG 719
 QY |||||
 4391 GGTCTTTTAT--CCTTGTCTGAGGAAACCATGACGGAAGGAAAGAACTGTGTTTAAAG 4448
 Db |||||
 720 GATTTTGTCTTGTGAGGAAACCATGAGGAAAGGAAAGGAACTGTGTTTAAAG 779
 QY |||||
 4449 GTCCACTTTTAAACTATATTTATTTATGTAG 4480
 Db |||||
 780 TCCACTTTTAAAGTATTAATTTTGGAG 811

RESULT 4
 BG568684 758 bp mRNA linear EST 10-APR-2001
 LOCUS 602587705F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4716384 5',
 DEFINITION mRNA sequence.
 ACCESSION BG568684
 VERSION BG568684.1 GI:13576337
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 758)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1564 row: j column: 01
High quality sequence stop: 753.

FEATURES		Location/Qualifiers	
Source		1.. 758	
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4716384"		
	/lab_host="DH10B (T1 phage-resistant)"		
	/clone_lib="NIH MGC 76"		
	/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGCGGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.85 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."		

ORIGIN

	Query Match	13.6%;	Score 662.4;	DB 4;	Length 758;
	Best Local Similarity	98.3%;	Pred. No. 2.6e-104;		
	Matches	743;	Conservative	0;	Mismatches 6;
				Indels	7;
				Gaps	7;
QY	4028	GATGAATGGAAAGGGGT	CATGGTACTCAATGAGGAAGATGAGTATGAGATCAGGCCCTT	4087	
Db	1	GATGAATGGAAAGGGGT	CATGGTACTCAATGAGGAAGATGAGTATGAGATCAGGCCCTT	60	
QY	4088	CTTCCACAGCAATAGTCCCACTACCTAGATTTTGTCTCTCTGTATGTGACAAATTT	4147		
Db	61	CTTCCACAGCAATAGTCCCACTACCTAGATTTTGTCTCTCTGTATGTGACAAATTT	119		
QY	4148	TTGTACATTATGTTATTTGGAATTTCTTTT	CATACATTATATTCCTCTTAAACTCTCAAGC	4207	
Db	120	TTGTACATTATGTTATTTGGAATTTCTTTT	CATACATTATATTCCTCTTAAACTCTCAAGC	179	
QY	4208	AGACGTGAGTGTGACTTTTGTGAAAAAGTATAGATAAATTACATTAATAATGACATGA	4267		
Db	180	AGACGTGAGTGTGACTTTTGTGAAAAAGTATAGATAAATTACATTAATAATGACATGA	239		
QY	4268	TTTTCTTTTGTTTTCTTCAATTTCTTTGTGCTCACCCCAAGAAAGTAACAAAAAGTATAGTTTGT	4327		
Db	240	TTTTCTTTTGTTTTCTTCAATTTCTTTGTGCTCACCCCAAGAAAGTAACAAAAAGTATAGTTTGT	297		
QY	4328	ACAGAGTTGGTTGTCATAATTTT	CAGTTCTAGTTGATTCGAGAAATTTTCAATAAGCAAG	4387	
Db	298	ACAGAGTTGGTTGTCATAATTTT	CAGTTCTAGTTGATTCGAGAAATTTTCAATAAGCAAG	357	
QY	4388	AGGGGTCTTTTATCTCTTGTCTGAGAAAAACCATCAGGAAGGAAAAAATCATGTTTAAA	4447		
Db	358	AGGGGTCTTTTATCTCTTGTCTGAGAAAAACCATCAGGAAGGAAAAAATCATGTTTAAA	417		
QY	4448	AGTCCACTTTTAAAAACTATATTTATTTATGTAGGATCTGTCAAAAGAAAACTTCCAAAAG	4507		
Db	418	AGTCCACTTTTAAAAACTATATTTATTTATGTAGGATCTGTCAAAAGAAAACTTCCAAAAG	477		
QY	4508	ATTTATTAATTTAAACCGACCTCTGTGTCAATAAGTTAATGTTTCTTGTGTTTGTAAATCCA	4567		
Db	478	ATTTATTAATTTAAACCGACCTCTGTGTCAATAAGTTAATGTTTCTTGTGTTTGTAAATCCA	536		
QY	4568	CACATTTCAATGAGTTAGGCTTTTGCACTTTGTAGGAAGGAGAGCGTTCAACAACCTCAAT	4627		
Db	537	CACATTTCAATGAGTTAGGCTTTTGCACTTTGTAGGAAGGAGAGCGTTCAACAACCTCAAT	596		

Institute (NEI).
TAG_TISSUE=human fetal eye
TAG_LIB=UI-E-E01
TAG_SEQ=CGCGTATACC"

ORIGIN

Query Match 13.4%; Score 655.2; DB 4; Length 679;
Best Local Similarity 99.4%; Pred. No. 4.7e-103;
Matches 668; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Other_GSSs: RPCI-11-265R19.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pietre de Jong
(pietre@dejong.med.buffalo.edu). Clones may be purchased from
BACpac Resources (<http://bacpac.med.buffalo.edu/ordering>) or
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.htm
Seq primer: SP6
Class: BAC ends.

FEATURES	SOURCE
1. Feature 1	Source 1
2. Feature 2	Source 2
3. Feature 3	Source 3
4. Feature 4	Source 4
5. Feature 5	Source 5
6. Feature 6	Source 6
7. Feature 7	Source 7
8. Feature 8	Source 8
9. Feature 9	Source 9
10. Feature 10	Source 10
11. Feature 11	Source 11
12. Feature 12	Source 12
13. Feature 13	Source 13
14. Feature 14	Source 14
15. Feature 15	Source 15
16. Feature 16	Source 16
17. Feature 17	Source 17
18. Feature 18	Source 18
19. Feature 19	Source 19
20. Feature 20	Source 20
21. Feature 21	Source 21
22. Feature 22	Source 22
23. Feature 23	Source 23
24. Feature 24	Source 24
25. Feature 25	Source 25
26. Feature 26	Source 26
27. Feature 27	Source 27
28. Feature 28	Source 28
29. Feature 29	Source 29
30. Feature 30	Source 30
31. Feature 31	Source 31
32. Feature 32	Source 32
33. Feature 33	Source 33
34. Feature 34	Source 34
35. Feature 35	Source 35
36. Feature 36	Source 36
37. Feature 37	Source 37
38. Feature 38	Source 38
39. Feature 39	Source 39
40. Feature 40	Source 40
41. Feature 41	Source 41
42. Feature 42	Source 42
43. Feature 43	Source 43
44. Feature 44	Source 44
45. Feature 45	Source 45
46. Feature 46	Source 46
47. Feature 47	Source 47
48. Feature 48	Source 48
49. Feature 49	Source 49
50. Feature 50	Source 50
51. Feature 51	Source 51
52. Feature 52	Source 52
53. Feature 53	Source 53
54. Feature 54	Source 54
55. Feature 55	Source 55
56. Feature 56	Source 56
57. Feature 57	Source 57
58. Feature 58	Source 58
59. Feature 59	Source 59
60. Feature 60	Source 60
61. Feature 61	Source 61
62. Feature 62	Source 62
63. Feature 63	Source 63
64. Feature 64	Source 64
65. Feature 65	Source 65
66. Feature 66	Source 66
67. Feature 67	Source 67
68. Feature 68	Source 68
69. Feature 69	Source 69
70. Feature 70	Source 70
71. Feature 71	Source 71
72. Feature 72	Source 72
73. Feature 73	Source 73
74. Feature 74	Source 74
75. Feature 75	Source 75
76. Feature 76	Source 76
77. Feature 77	Source 77
78. Feature 78	Source 78
79. Feature 79	Source 79
80. Feature 80	Source 80
81. Feature 81	Source 81
82. Feature 82	Source 82
83. Feature 83	Source 83
84. Feature 84	Source 84
85. Feature 85	Source 85
86. Feature 86	Source 86
87. Feature 87	Source 87
88. Feature 88	Source 88
89. Feature 89	Source 89
90. Feature 90	Source 90
91. Feature 91	Source 91
92. Feature 92	Source 92
93. Feature 93	Source 93
94. Feature 94	Source 94
95. Feature 95	Source 95
96. Feature 96	Source 96
97. Feature 97	Source 97
98. Feature 98	Source 98
99. Feature 99	Source 99
100. Feature 100	Source 100

```

FEATURES
source
    class: bac_end.
    location/Qualifiers
    1. .673
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="GDB:7601490"
    /db_xref="taxon:9606"
    /clone="RPC1-11-265E19"
    /sex="Male"
    /cell_type="Lymphocytes"
    /clone_lib="RPC1-11"
    /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
    RPC11 Human Male BAC Library"

```

```

QY 2947 TTTATTTTGTCTTTCTTTAGTGAATATTTGGCTTGGAAATGATAAAATAGCCAGCTTA 3006
Db 540 TTTATTTTGTATTTCTTTTAGTGAATATTTGGCTTGGAAATGATAAAATAGCCAGCTTA 599
QY 3007 CCAGATGGGACCCACAGAACTTTTGTATAGAAATGGAGGAGTGGAAAGGAGACAAAGTAA 3066
Db 600 CCAGATGGGACCCACAGAACTTTTGTATAGAAATGGAGGAGTGGAAAGGAGACAAAGTAA 659
QY 3067 AGGCTCACTATGGA 3080
Db 660 AGGCTCACTATGGA 673

RESULT 7
CR594769 1891 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DM011Y115 of Fetal liver of Homo sapiens
(human).
ACCESSION CR594769
VERSION CR594769.1 GI:50475576
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1891)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 1891)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1891
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM011Y115"
/tissue_type="Fetal liver"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 13.2%; Score 644.4; DB 3; Length 1891;
Best Local Similarity 99.7%; Pred. No. 2,9e-101;
Matches 656; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3869 CAGGTTACATCAGATCCAGAAAACAGTGTCTTAAGAGACGGTGGTGGATGGTGA 3928
Db 1235 CTGGTTACATCAGATCCAGAAAACAGTGTCTTAAGAGACGGTGGTGGATGGTGA 1294
QY 3929 TAATAGATGTCATGACGCAATCCAAACGGCAGATCTACTGGGTGGACAGTACACCTG 3988
Db 1295 TAATAGATGTCATGACGCAATCCAAACGGCAGATCTACTGGGTGGACAGTACACCTG 1354
QY 3989 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGGATGAATGGAAGGGTCAATG 4048
Db 1355 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGGATGAATGGAAGGGTCAATG 1414
QY 4049 GTACTCAATGAGAGATGAGTATGAGATCAGGCCCTTCTTCCACACCAATAGTCCCC 4108
Db 1415 GTACTCAATGAGAGATGAGTATGAGATCAGGCCCTTCTTCCACACCAATAGTCCCC 1474

```

```

QY 4109 AATACGTAGATTTTGTCTCTTCTGTATGTGACAACTTTTGTGATGATGTTATGGAA 4168
Db 1475 AATACGTAGATTTTGTCTCTTCTGTATGTGACAACTTTTGTGATGATGTTATGGAA 1534
QY 4169 TTTCTTTTCATACATATATATCTCTTAAACTCTCAAGCAGACGCTGAGTGTGACTTTTGTG 4228
Db 1535 TTTCTTTTCATACATATATATCTCTTAAACTCTCAAGCAGACGCTGAGTGTGACTTTTGTG 1594
QY 4229 AAAAAAGTATAGGATAAAATACATTTAAATAGCACATGATTTCTTTTGTCTTCTCAT 4288
Db 1595 AAAAAAGTATAGGATAAAATACATTTAAATAGCACATGATTTCTTTTGTCTTCTCAT 1654
QY 4289 TCTCTTCTCACCACCAAGAGTAACAAAAGTATAGTTTTCAGACAGATGGTGTTCATTAAT 4348
Db 1655 TCTCTTCTCA-CCAAGAGTAACAAAAGTATAGTTTTCAGACAGATGGTGTTCATTAAT 1713
QY 4349 TCAGTTCTAGTTGATTCGGAGAATTTTCAATAAGGAAGGGGTCTTTTATCTTGTGCG 4408
Db 1714 TCAGTTCTAGTTGATTCGGAGAATTTTCAATAAGGAAGGGGTCTTTTATCCTTGTGCG 1773
QY 4409 TAGGAAAACCATGACGGAAAGGAAAACTGATGTTTAAAGTCCACITTTTAAACTATAT 4468
Db 1774 TAGGAAAACCATGACGGAAAGGAAAACTGATGTTTAAAGTCCACITTTTAAACTATAT 1833
QY 4469 TTATTTATGTAGGATCTGTCAAAGAAAACCTTCCAAAAGATTTATTAATTAACACAGA 4526
Db 1834 TTAATTATGTAGGATCTGTCAAAGAAAACCTTCCAAAAGATTTATTAATTAACACAGA 1891

RESULT 8
CR596720 1577 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DM009Y104 of Fetal liver of Homo sapiens
(human).
ACCESSION CR596720
VERSION CR596720.1 GI:50477527
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1577)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 1577)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1577
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM009Y104"
/tissue_type="Fetal liver"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 13.2%; Score 643.4; DB 3; Length 1577;
Best Local Similarity 99.7%; Pred. No. 4,5e-101;
Matches 655; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES	SOURCE
Location/Qualifiers	1, 1883
organism="Homo sapiens"	
mol type="mRNA"	
db_xref="taxon:9606"	
clone="CS0DM009YCl8"	
tissue type="Fetal liver"	
plasmid="pCMVSPORT 2"	

ORIGIN

Query Match 13.2%; Score 642.4; DB 3; Length 1883;
Best Local Similarity 99.7%; Pred. No. 6.5e-101;
Matches 654; Conservative 0; Mismatches 1; Indels 1;

3869	QY	CAGGTTAA	CATCAGATCC	CAGAAACAA	CAGTGTCT	TAAAGAA	GACGGTGGT	CGATGGTGA	3928
1229	Db	CTGGTTAA	CATCAGATCC	CAGAAACAA	CAGTGTCT	TAAAGAA	GACGGTGGT	CGATGGTGA	1288
3929	QY	TAATAGAT	GTCAATCG	CAGCCAA	TCTACT	CGGGTGG	ACAGTAC	ACACCTG	3988
1289	Db	TAATAGAT	GTCAATCG	CAGCCAA	TCTACT	CGGGTGG	ACAGTAC	ACACCTG	1348
3989	QY	GGA	CATGG	CAAA	GCATGG	CA	GATGGT	TAGTATGG	4048
1349	Db	GGA	CATGG	CAAA	GCATGG	CA	GATGGT	TAGTATGG	1408
4049	QY	GTA	CTCAAT	CAGAA	AGATGA	GTGA	ATCAGG	CCCTCTCC	4108
1409	Db	GTA	CTCAAT	CAGAA	AGATGA	GTGA	ATCAGG	CCCTCTCC	1468
4109	QY	AA	TACGTAG	ATTTT	GCTCTCT	TGATGT	GACAA	CAATTTTG	4168
1469	Db	AA	TACGTAG	ATTTT	GCTCTCT	TGATGT	GACAA	CAATTTTG	1528
4169	QY	TTTTCTTT	CATACAT	TATATCT	CTAA	AACTCT	CAAG	CAGACGTG	4228
1529	Db	TTTTCTTT	CATACAT	TATATCT	CTAA	AACTCT	CAAG	CAGACGTG	1588
4229	QY	AAAAAG	TATAG	ATAAA	TTTAC	ATAAAA	TAG	CAATGATTTT	4288
1589	Db	AAAAAG	TATAG	ATAAA	TTTAC	ATAAAA	TAG	CAATGATTTT	1648
4289	QY	TCTCTT	GCTCA	CCCA	AGAA	GTAA	CAAA	GTATAGTTT	4348
1649	Db	TCTCTT	GCTCA	CCCA	AGAA	GTAA	CAAA	GTATAGTTT	1707
4349	QY	TCAG	TTCTAG	TTGAT	TCGAG	AAATTTT	CAAA	TAAAGAGAGGGG	4408
1708	Db	TCAG	TTCTAG	TTGAT	TCGAG	AAATTTT	CAAA	TAAAGAGAGGGG	1767
4409	QY	TAGG	AAAA	CCAT	GA	CGAA	AGAAA	AACTGATGTTT	4468
1768	Db	TAGG	AAAA	CCAT	GA	CGAA	AGAAA	AACTGATGTTT	1827
4469	QY	TTAT	TTAT	GTAG	ATCT	GTCA	AA	AGAAA	4524
1828	Db	TTAT	TTAT	GTAG	ATCT	GTCA	AA	AGAAA	1883

RESULT 10	
CR6111717	
LOCUS	CR6111717
DEFINITION	1889 bp mRNA linear HTC 21-JUL-2004 full-length cDNA clone CSODM011B17 of Fetal liver of Homo sapiens (human).
ACCESSION	CR6111717
VERSION	CR6111717.1 GI:50492524
KEYWORDS	HTC; CNSLT_cDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

3969	Qy	CAGTTTAACTACAGATCCAGAAACACAGTGTCTTAAAGAAACAGCGTGGTGGATGGTGGTA	3928
922	Db	CTGTTTAACTACAGATCCAGAAACACAGTGTCTTAAAGAAACAGCGTGGTGGATGGTGGTA	981
3929	Qy	TAATAGATGTCACTCAGCCCAATCCAAACGGCAGATACTACTGGGGTGGACAGTACACCTG	3988
982	Db	TAATAGATGTCACTCAGCCCAATCCAAACGGCAGATACTACTGGGGTGGACAGTACACCTG	1041
3989	Qy	GGACATGGCAAAAGCATGGCACAGATGATGGTGTAGTATGGATGAATTTGGAAGGGGTCAATG	4048
1042	Db	GGACATGGCAAAAGCATGGCACAGATGATGGTGTAGTATGGATGAATTTGGAAGGGGTCAATG	1101
4049	Qy	GTACTCAATGAGGAGATGATGATGAAGATCAGGCCCTTCTTCCACAGCAATAGTATCCCC	4108
1102	Db	GTACTCAATGAGGAGATGATGATGAAGATCAGGCCCTTCTTCCACAGCAATAGTATCCCC	1161
4109	Qy	AATACGTAGATTTTGTCTCTCTGTATGTGCACAACATTTTTGTACATTATGTTATTGGAA	4168
1162	Db	AATACGTAGATTTTGTCTCTCTGTATGTGCACAACATTTTTGTACATTATGTTATTGGAA	1221
4169	Qy	TTTTCTTTTCATACTATATATTCCTCTTAAACTCTCAAGCAGACGTGAGTGTGACTTTTGTG	4228
1222	Db	TTTTCTTTTCATACTATATATTCCTCTTAAACTCTCAAGCAGACGTGAGTGTGACTTTTGTG	1281
4229	Qy	AAAAAGATATAGATATAATTACATTAAAAATAGCACATGATTTTCTTTGTGTTTCTTCATT	4288
1282	Db	AAAAAGATATAGATATAATTACATTAAAAATAGCACATGATTTTCTTTGTGTTTCTTCATT	1341
4289	Qy	TCCTTTGCTCACCAAGAAGTAAACAAAGTATATAGTTTTCACAGAGTTGGTGTTCATAATT	4348
1342	Db	TCCTTTGCTCA - CCAAGAAGTAAACAAAGTATATAGTTTTCACAGAGTTGGTGTTCATAATT	1400
4349	Qy	TCAGTTTCTAGTTCAATTGCGAGAAATTTTCAAATGAAGAAAGGGGTCTTTTATCTCTGTGCG	4408
1401	Db	TCAGTTTCTAGTTCAATTGCGAGAAATTTTCAAATGAAGAAAGGGGTCTTTTATCTCTGTGCG	1460
4409	Qy	TAGGAAAAACCATCAGCGAAGGAAAAACTGATCTTTTAAAGTCCACTTTTAAACCTATAT	4468
1461	Db	TAGGAAAAACCATCAGCGAAGGAAAAACTGATCTTTTAAAGTCCACTTTTAAACCTATAT	1520
4469	Qy	TTATTTTATGAGATCTGTCAAGAAACCTTCAAAAAGATTTTATTAATTAACACAG	4525
1521	Db	TTATTTTATGAGATCTGTCAAGAAACCTTCAAAAAGATTTTATTAATTAACACAG	1577

RESULT 9	
CR605676	
LOCUS	1883 bp mRNA linear HTC 21-JUL-2004
DEFINITION	full-length cDNA clone CS0DM009XG18 of Fetal liver of Homo sapiens (human)
ACCESSION	CR605676
VERSION	CR605676.1 GI:50486483
KEYWORDS	HTC; CNSLT_cDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1. (bases 1 to 1883)
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
REMARK	Contact : Peng Liang Email : pliang@lifetech.com

REFERENCE	faraday Avenue
AUTHORS	2 (bases 1 to 1883)
TITLE	Genoscope.
JOURNAL	Direct Submission
COMMENT	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr) - Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo (5') primer. Bino...

```

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1888)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 1888)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
18c strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
    source
        Location/Qualifiers
            1..1888
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CSODM011YB17"
                /tissue_type="Fetal Liver"
                /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      13.1%; Score 641.4; DB 3; Length 1888;
Best Local Similarity 99.7%; Pred. No. 9.7e-101;
Matches 653; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 3869 CAGGTTAATCATCAGATCCAGAAACAGTGTCTTAAAGAACGCGTGGTGGTGGTGA 3928
DB 1235 CTGGTTAATCATCAGATCCAGAAACAGTGTCTTAAAGAACGCGTGGTGGTGGTGA 1294
QY 3929 TAATAGATGTCTATGAGCAATCCAAACGCGAGATCTACTGGGTTGACAGTACACCTG 3988
DB 1295 TAATAGATGTCTATGAGCAATCCAAACGCGAGATCTACTGGGTTGACAGTACACCTG 1354
QY 3989 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGATGATGATGATGATGATGATG 4048
DB 1355 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGATGATGATGATGATGATGATG 1414
QY 4049 GTACTCAATGAGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4108
DB 1415 GTACTCAATGAGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1474
QY 4109 AATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4168
DB 1475 AATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1534
QY 4169 TTTCTTTTATCATATATATTTCTTAAACTCTCAAGCAGACGCTGAGTGGTGGTGGTGGT 4228
DB 1535 TTTCTTTTATCATATATATTTCTTAAACTCTCAAGCAGACGCTGAGTGGTGGTGGTGGT 1594
QY 4229 AAAAAAGTATAGGATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4288
DB 1595 AAAAAAGTATAGGATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1654
QY 4289 TCTCTTGTCTACCCCAAGAGTAAACAAAGTATAGTATGATGATGATGATGATGATGATG 4348
DB 1655 TCTCTTGTCTACCCCAAGAGTAAACAAAGTATAGTATGATGATGATGATGATGATGATG 1713
QY 4349 TCAGTTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4408
DB 1714 TCAGTTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1773
QY 4409 TAGGAAAAACATGACCGGAAGGAAACATGATGATGATGATGATGATGATGATGATGATG 4468
DB 1774 TAGGAAAAACATGACCGGAAGGAAACATGATGATGATGATGATGATGATGATGATGATG 1833

```

```

4469 TTATTATGTAGGATCTGTCAAGAAACACTTCCAAAGAGATTTATTATTAAATTAACC 4523
1834 TTATTATGTAGGATCTGTCAAGAAACACTTCCAAAGAGATTTATTATTAAATTAACC 1888

```

```

RESULT 11
BX448857/c
LOCUS
DEFINITION
    BX448857 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
    CSODM012YE15 3-PRIME, mRNA sequence.
ACCESSION
    BX448857
VERSION
    GI:47063272
KEYWORDS
    EST.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
    1 (bases 1 to 1014)
    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
    Full-length cDNA libraries and normalization
    Unpublished (2001)
    On May 22, 2003 this sequence version replaced gi:31025989.
    Contact: Genoscope
    Genoscope - Centre National de Sequencage
    BP 191 91006 Evry cedex - France
    Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
    1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
    end enriched, double-strand cDNA was digested with Not I and cloned
    into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
    was not normalized. Library was constructed by Life Technologies, a
    division of Invitrogen.
    This sequence belongs to sequence cluster 2627.f
    For more information about this cluster, see
    http://www.genoscope.cns.fr/cdna?s=CS0AM012AC08NP1&c=2627.f.
FEATURES
    source
        Location/Qualifiers
            1..1014
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CSODM012YE15"
                /tissue_type="FETAL LIVER"
                /dev_stage="fetal"
                /clone_lib="Homo sapiens FETAL LIVER"
                /note="Organ: liver; vector: pCMVSPORT_6; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-strand cDNA was digested with Not I and
                cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                vector. Library was not normalized."
ORIGIN
Query Match      13.1%; Score 640.4; DB 5; Length 1014;
Best Local Similarity 99.1%; Pred. No. 1.6e-100;
Matches 650; Conservative 4; Mismatches 1; Indels 1; Gaps 1;
QY 3869 CAGGTTAATCATCAGATCCAGAAACAGTGTCTTAAAGAACGCGTGGTGGTGGTGA 3928
DB 655 CTGGTTAATCATCAGATCCAGAAACAGTGTCTTAAAGAACGCGTGGTGGTGGTGGTGA 596
QY 3929 TAATAGATGTCTATGAGCAATCCAAACGCGAGATCTACTGGGTTGACAGTACACCTG 3988
DB 595 TAATAGATGTCTATGAGCAATCCAAACGCGAGATCTACTGGGTTGACAGTACACCTG 536
QY 3989 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGATGATGATGATGATGATGATG 4048
DB 535 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGATGATGATGATGATGATGATG 476
QY 4049 GTACTCAATGAGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4108
DB 475 GTACTCAATGAGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 416
QY 4109 AATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4168
DB 415 AATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 356

```


QY 4169 TTTCTTTTCATATATATCTCTAAACTCTCAAGCAGACGTGAGTGTGACTTTTTCG 4228
 |||||
 Db 355 TTTTCTTTCATATATATCTCTAAACTCTCAAGCAGACGTGAGTGTGACTTTTTCG 296
 |||||
 QY 4229 AAAAAAGTAGATATAATTTACATTAATAATAGCAGATGATTTTCTTTTGTCTTCATT 4288
 |||||
 Db 295 AAAAAAGTAGATATAATTTACATTAATAATAGCAGATGATTTTCTTTTGTCTTCATT 236
 |||||
 QY 4289 TCTCTTGTCTCACCACCAAGAGTAACTCAAAAGTATAGTTTTCAGAGATTTGGTGTTCATAAT 4348
 |||||
 Db 235 TCTCTTGTCTCA - CCAAGAGTAACTCAAAAGTATAGTTTTCAGAGATTTGGTGTTCATAAT 177
 |||||
 QY 4349 TCAGTTTCTAGTTGAGTGTGAGATTTTCAAAATAGGAAGAGGGGTCTTTTATCTTTGTGCG 4408
 |||||
 Db 176 TCAGTTTCTAGTTGAGTGTGAGATTTTCAAAATAGGAAGAGGGGTCTTTTATCTTTGTGCG 117
 |||||
 QY 4409 TAGGAAGACCATGACGGAAAGGAAAGTGTGATTTTAAAGTCCATTTTAAAGTATAT 4468
 |||||
 Db 116 TAGGAAGACCATGACGGAAAGGAAAGTGTGATTTTAAAGTCCATTTTAAAGTATAT 57
 |||||
 QY 4469 TTATTTATGTAGTATCTGTCAAGAAAGTCTTCAAAAGAGATTTAAATTAACCA 4524
 |||||
 Db 56 TTATTTATGTAGTATCTGTCAAGAAAGTCTTCAAAAGAGATTTAAATTAACCA 1

RESULT 12
 CR596719
 LOCUS full-length cDNA clone CS0DM009Y103 of Fetal liver of Homo sapiens (human).
 DEFINITION
 ACCESSION CR596719.1 GI:50477526
 VERSION
 KEYWORDS HTC; CNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1886)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 JOURNAL
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 Genoscope.
 2 (bases 1 to 1886)
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 Location/Qualifiers
 1..1886
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DM009Y103"
 /tissue_type="Fetal liver"
 /plasmid="pCMVSPORT_6"

ORIGIN
 Query Match 13.1%; Score 639.4; DB 3; Length 1886;
 Best Local Similarity 99.7%; Pred. No. 2.2e-100;
 Matches 651; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 3869 CAGGTTACATCAGATCCAGAAAACAGTGTCTAAAGAACGCGTGGTGGTGGTA 3928
 |||||
 Db 1235 CTGGTTAAACATCAGATCCAGAAAACAGTGTCTAAAGAACGCGTGGTGGTGGTA 1294

QY 3929 TAATAGATGTCATGCGAGCCATCTCAAAAGCAGATACTACTGGGGTGGACAGTACACCTG 3988
 |||||
 Db 1295 TAATAGATGTCATGCGAGCCATCTCAAAAGCAGATACTACTGGGGTGGACAGTACACCTG 1354
 |||||
 QY 3989 GGACATGCGAAAAGCAGTGGCAGACAGATGATGTTGATGATCGATGAATTTGGAAGGGGTCAATG 4048
 |||||
 Db 1355 GGACATGCGAAAAGCAGTGGCAGACAGATGATGTTGATGATCGATGAATTTGGAAGGGGTCAATG 1414
 |||||
 QY 4049 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCTC 4108
 |||||
 Db 1415 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCTC 1474
 |||||
 QY 4109 AATACGTAGATTTTGTCTCTTCTGTATGTGACAAAGATTTTGTACATATGTTATTTGGAA 4168
 |||||
 Db 1475 AATACGTAGATTTTGTCTCTTCTGTATGTGACAAAGATTTTGTACATATGTTATTTGGAA 1534
 |||||
 QY 4169 TTTTCTTTTCATATATATCTCTTAAAACTCTCAAGCAGACGTGAGTGTGACTTTTTCG 4228
 |||||
 Db 1535 TTTTCTTTTCATATATATCTCTTAAAACTCTCAAGCAGACGTGAGTGTGACTTTTTCG 1594
 |||||
 QY 4229 AAAAAAGTAGATAAATTTACATTAATAATAGCAGATGATTTTCTTTTGTCTTCATT 4288
 |||||
 Db 1595 AAAAAAGTAGATAAATTTACATTAATAATAGCAGATGATTTTCTTTTGTCTTCATT 1654
 |||||
 QY 4289 TCTCTTGTCTCACCACCAAGAGTAACTCAAAAGTATAGTTTTCAGAGATTTGGTGTTCATAAT 4348
 |||||
 Db 1655 TCTCTTGTCTCA - CCAAGAGTAACTCAAAAGTATAGTTTTCAGAGATTTGGTGTTCATAAT 1713
 |||||
 QY 4349 TCAGTTTCTAGTTGAGTGTGAGATTTTCAAAATAGGAAGAGGGGTCTTTTATCTTTGTGCG 4408
 |||||
 Db 1714 TCAGTTTCTAGTTGAGTGTGAGATTTTCAAAATAGGAAGAGGGGTCTTTTATCTTTGTGCG 1773
 |||||
 QY 4409 TAGGAAGACCATGACGGAAAGGAAAGTGTGATTTTAAAGTCCATTTTAAAGTATAT 4468
 |||||
 Db 1774 TAGGAAGACCATGACGGAAAGGAAAGTGTGATTTTAAAGTCCATTTTAAAGTATAT 1833
 |||||
 QY 4469 TTATTTATGTAGTATCTGTCAAGAAAGTCTTCAAAAGAGATTTAAATTAACCA 4521
 |||||
 Db 1834 TTATTTATGTAGTATCTGTCAAGAAAGTCTTCAAAAGAGATTTAAATTAACCA 1886
 |||||

RESULT 13
 BX427240/c
 LOCUS BX427240 Homo sapiens FETAL LIVER Homo sapiens cDNA clone CS0DM009Y104 3-PRIME, mRNA linear EST 04-MAY-2004
 DEFINITION
 ACCESSION BX427240
 VERSION
 KEYWORDS EST.
 SOURCE BX427240.1 GI:30772645
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 997)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 This sequence belongs to sequence cluster 2627.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?e=CS0AM009BE02NP1&c=2627.f.
 FEATURES
 Location/Qualifiers
 1..997
 /organism="Homo sapiens"

```
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="CS0DM009Y104"  
/tissue_type="FETAL LIVER"  
/dev_stage="fetal"  
/clone_lib="Homo sapiens FETAL LIVER"  
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-strand cDNA was digested with Not I and  
cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
vector. Library was not normalized."  
  
ORIGIN  
Query Match 13.1%; Score 637.8; DB 5; Length 997;  
Best Local Similarity 98.3%; Pred. No. 4.5e-100;  
Matches 646; Conservative 7; Mismatches 3; Indels 1; Gaps 1;  
  
QY 3869 CAGGTTAAACATCAGATCCAGAAACACAGTGTCTTAAAGAACGCGTGTGGATGGTGA 3928  
DB 656 CTGGTTAAACATCAGATCCWAAAAACAGTGTCTTAAAGAACGCGTGTGGATGGTGA 597  
  
QY 3929 TAATAGATGTCTATGCGAGCAATCCAAACGCGAGATCTACTGGGTGGACAGTACACCTG 3988  
DB 596 TAATAGATGTCTATGCGAGCAATCCAAACGCGAGATCTACTGGGTGGACAGTACACCTG 537  
  
QY 3989 GGACATGCCAAAGCATGGCACAGATGATGGTGTAGTATGGATGAATGGAAGGGGTCTATG 4048  
DB 536 GGACATGCCAAAGCATGGCACAGATGATGGTGTAGTATGGATGAATGGAAGGGGTCTATG 477  
  
QY 4049 GTACTCAATAGGAAAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 4108  
DB 476 GTACTCAATAGGAAAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 417  
  
QY 4109 AATAGTAGATTTTGGCTTCTCTGTATGTGACAAATTTTGTACATTTATTTGGAA 4168  
DB 416 AATAGTAGATTTTGGCTTCTCTGTATGTGACAAATTTTGTACATTTATTTGGAA 357  
  
QY 4169 TTTTCTTTCATACATTTATTTCTCTTAAACCTCTCAAGCAGACGCTGAGTGTGACATTTTGG 4228  
DB 356 TTTTCTTTCATACATTTATTTCTCTTAAACCTCTCAAGCAGACGCTGAGTGTGACATTTTGG 297  
  
QY 4229 AAAAAAGTAGGATAAATTAATCAATTAATAAGCATGATTTCTTTTGTCTTCTCAT 4288  
DB 296 AAAAAAGTAGGATAAATTAATCAATTAATAAGCATGATTTCTTTTGTCTTCTCAT 237  
  
QY 4289 TCTCTTGTCTCACCAGAAAGTAACAAAGTATAGTTTTCACAGATGCTGTGTTTCATATT 4348  
DB 236 TCTCTTGTCTCA-CCAAGAAAGTAACAAAGTATAGTTTTCACAGATGCTGTGTTTCATATT 178  
  
QY 4349 TCAGTTCTAGTTGATGCGAGAAATTTCAATAAGGAAGAGGGGTCTTTTATCCTTGTGCG 4408  
DB 177 TCAGTTCTAGTTGATGCGAGAAATTTCAATAAGGAAGAGGGGTCTTTTATCCTTGTGCG 118  
  
QY 4409 TAGGAAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACATATAT 4468  
DB 117 TAGGAAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACATATAT 58  
  
QY 4469 TTATTATGATAGATCTGTCAAGAAACCTTCCAAAGAGATTTATTAATTAACACG 4525  
DB 57 TTATTATGATAGATCTGTCAAGAAACCTTCCAAAGAGATTTVKNKAADWAACACG 1  
  
RESULT 14  
CR626133  
LOCUS  
DEFINITION  
full-length cDNA clone CS0DH008Y104 of T cells (Jurkat cell line)  
of Homo sapiens (human).  
ACCESSION  
CR626133  
VERSION  
CR626133.1 GI:50506940  
KEYWORDS  
HTC; CNSLT_cDNA.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1880)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1880)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: secref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. 1880
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH008Y104"
/tissue_type="T cells (Jurkat cell line)"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 13.0%; Score 634.4; DB 3; Length 1880;
Best Local Similarity 99.7%; Pred. No. 1.6e-99;
Matches 646; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3869 CAGGTTAAACATCAGATCCAGAAACACAGTGTCTTAAAGAACGCGTGTGGATGGTGA 3928
DB 1234 CTGGTTAAACATCAGATCCAGAAACACAGTGTCTTAAAGAACGCGTGTGGATGGTGA 1293

QY 3929 TAATAGATGTCTATGCGAGCAATCCAAACGCGAGATCTACTGGGTGGACAGTACACCTG 3988
DB 1294 TAATAGATGTCTATGCGAGCAATCCAAACGCGAGATCTACTGGGTGGACAGTACACCTG 1353

QY 3989 GGACATGCCAAAGCATGGCACAGATGATGGTGTAGTATGGATGAATTCGAAGGGGTCTATG 4048
DB 1354 GGACATGCCAAAGCATGGCACAGATGATGGTGTAGTATGGATGAATTCGAAGGGGTCTATG 1413

QY 4049 GTACTCAATAGGAAAGTATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 4108
DB 1414 GTACTCAATAGGAAAGTATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 1473

QY 4109 AATAGTAGATTTTGGCTTCTCTGTATGTGACAAATTTTGTGATCATTTATTTGGAA 4168
DB 1474 AATAGTAGATTTTGGCTTCTCTGTATGTGACAAATTTTGTGATCATTTATTTGGAA 1533

QY 4169 TTTTCTTTCATACATTTATTTCTCTTAAACCTCTCAAGCAGACGCTGAGTGTGACATTTTGG 4228
DB 1534 TTTTCTTTCATACATTTATTTCTCTTAAACCTCTCAAGCAGACGCTGAGTGTGACATTTTGG 1593

QY 4229 AAAAAAGTAGGATAAATTAATCAATTAATAAGCATGATTTTCTTTTCTTCTCAT 4288
DB 1594 AAAAAAGTAGGATAAATTAATCAATTAATAAGCATGATTTTCTTTTCTTCTCAT 1653

QY 4289 TCTCTTGTCTCACCAGAAAGTAACAAAGTATAGTTTTCACAGATGCTGTGTTTCATATT 4348
DB 1654 TCTCTTGTCTCA-CCAAGAAAGTAACAAAGTATAGTTTTCACAGATGCTGTGTTTCATATT 1712

QY 4349 TCAGTTCTAGTTGATGCGAGAAATTTTCAATAAGGAAGAGGGGTCTTTTATCCTTGTGCG 4408
DB 1713 TCAGTTCTAGTTGATGCGAGAAATTTTCAATAAGGAAGAGGGGTCTTTTATCCTTGTGCG 1772

QY 4409 TAGGAAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACATATAT 4468
DB 1773 TAGGAAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACATATAT 1832

QY 4469 TTATTATAGGATCTGTCAAGAAACTTCCAAAAGATTATTAA 4516
 Db ||||||| GTACTCAATGAGAGATGAGTATGAGATCAGGCCCTCTTCCACAGCAATAGTCCCC 430
 1833 TTATTATAGGATCTGTCAAGAAACTTCCAAAAGATTATTAA 1880
 |||||||

RESULT 15
 AW003933/c
 LOCUS
 DEFINITION
 similar to gb:M64983_rna6 FIBRINOGEN BETA CHAIN PRECURSOR (HUMAN);,
 mRNA sequence.

ACCESSION
 AW003933 804 bp mRNA linear EST 08-SEP-1999
 VERSION
 wq84e02.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2478746 3,
 similar to gb:M64983_rna6 FIBRINOGEN BETA CHAIN PRECURSOR (HUMAN);,
 mRNA sequence.

KEYWORDS
 EST.

SOURCE
 AW003933.1 GI:5850849
 Homo sapiens (human)

ORGANISM
 Homo sapiens

REFERENCE
 1 (bases 1 to 804)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

TITLE
 Unpublished (1997)

JOURNAL
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov

COMMENT
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 463.

FEATURES
 Location/Qualifiers
 1..804
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2478746"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP GC6"
 /note="Vector: pYT3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not 1; Site_2: Eco RI; Plasmid DNA
 from the normalized library NCI CGAP GC4 was prepared, and
 ss circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (clones
 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 13.0%; Score 634; DB 1; Length 804;
 Best Local Similarity 97.6%; Pred. No. 2.1e-99;
 Matches 654; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 QY 3869 CAGGTTAAACATCAGATCCAGAAACAGTGTCTAAAGAACGCGTGGTGGTGGTA 3928
 Db ||||||| CAGGTTAAACATCAGATCCAGAAACAGTGTCTAAAGAACGCGTGGTGGTGGTA 610
 669 CTGGTTAAACATCAGATCCAGAAACAGTGTCTAAAGAACGCGTGGTGGTGGTA 610
 |||||||

QY 3929 TAATAGATGTCATGACGCAATCCAAACGGCAGATACCTAGTGGGTGGACAGTACACCTG 3988
 Db ||||||| TAATAGATGTCATGACGCAATCCAAACGGCAGATACCTAGTGGGTGGACAGTACACCTG 550
 609 TAATAGATGTCATGACGCAATCCAAACGGCAGATACCTAGTGGGTGGACAGTACACCTG 550
 |||||||

QY 3989 GGACATGGCAAGCATGGCAGAGATGGTGTAGTATGCGAATTCGAAGGGGTGATG 4048
 Db ||||||| GGACATGGCAAGCATGGCAGAGATGGTGTAGTATGCGAATTCGAAGGGGTGATG 490
 549 GGACATGGCAAGCATGGCAGAGATGGTGTAGTATGCGAATTCGAAGGGGTGATG 490
 |||||||

QY 4049 GTACTCAATGAGGAGATGAGTATGAGATCAGGCCCTCTTCCACAGCAATAGTCCCC 4108
 ||||||| GTACTCAATGAGGAGATGAGTATGAGATCAGGCCCTCTTCCACAGCAATAGTCCCC 430

Search completed: November 23, 2004, 12:26:29
 Job time : 14403.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 18:22:12 ; Search time 155 Seconds
(without alignments)
1136.362 Million cell updates/sec

Title: US-10-017-724-6
Perfect score: 2676
Sequence: 1 MKRMVSWFPHKLTWKHL.....GSWYSMRKSMKIRFFPQQ 491

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2676	100.0	491	6	Aae36413 Human FBG
2	2676	100.0	491	7	Adp65229 Human fib
3	2676	100.0	491	7	Adp65300 Human fib
4	2671	99.8	491	2	Aar82243 Human fib
5	2671	99.8	491	4	Aam78492 Human pro
6	2651	99.1	495	4	Aam79475 Human pro
7	2651	99.1	495	4	Aam79476 Human pro
8	2651	99.1	495	4	Aam79477 Human pro
9	2648	99.0	488	8	Adc76868 Human pro
10	2637	98.5	539	4	Aam78491 Human pro
11	2559	95.6	491	7	Adp69271 Human fib
12	2559	95.6	491	7	Adp69272 Human fib
13	2436	91.0	453	4	Aam78493 Human pro
14	2170	81.1	479	8	Adi82107 Fibrinoge
15	1072.5	40.1	489	5	Abu60918 Lung spec
16	1001	37.4	502	4	Aau31380 Novel hum
17	785	29.3	150	3	Aag00150 Human sec
18	630.5	23.6	437	8	Adk70498 Respirato
19	630.5	23.6	437	8	Adi61247 Human pro
20	630.5	23.6	453	7	Adi61247 Human pro
21	630.5	23.6	453	8	Adc45364 Human pro
22	629.5	23.5	453	8	Adc30579 Pancreas
23	629.5	23.5	453	2	Aar82246 Human fib
24	620.5	22.2	411	2	Aar82245 Human fib
25	610	22.8	141	4	Aar84650 Fibrinoge
				4	Aao10778 Human pol

ALIGNMENTS

RESULT 1
AAE36413
ID AAE36413 standard; protein; 491 AA.
XX
AC AAE36413;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human FBG reference protein (GI 11761631).

XX
KW Thrombospondin 2; THBS2; angiotensin converting enzyme; polymorphism;
ACE-1; beta-fibrinogen; FGB; peripheral vascular disease; ischaemia;
KW vascular disease; myocardial infarction; pulmonary embolism; stroke;
KW atherosclerosis; coronary artery disease; venous thromboembolism; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 478
FT /note= "This residue is changed to Lys due to single
FT nucleotide polymorphism (SNP)"
XX
PN WO2003020118-A2.
XX
PD 13-MAR-2003.
XX
PF 04-SEP-2002; 2002WO-US028113.
XX
PR 05-SEP-2001; 2001US-0317178P.
PR 16-OCT-2001; 2001US-0329958P.
PR 14-DEC-2001; 2001US-00017724.
XX
PA (VITI-) VITIVITY INC.
XX
PI McCarthy J;
XX
DR WPI; 2003-300816/29.
XX
PT Identifying polymorphisms in thrombospondin 2, angiotensin converting
PT enzyme and/or beta-fibrinogen genes in nucleic acid sample of subject, by
PT contacting the nucleic acid with a complementary probe or primer.
XX
PS Claim 5; Fig 6; 194pp; English.
XX
CC The invention relates to a method for determining the identity of one or
CC more allelic variants of a polymorphic region of a thrombospondin 2
CC (THBS2), angiotensin converting enzyme (ACE)-1 and/or beta-fibrinogen
CC (FGB) genes in a nucleic acid obtained from a subject. The method

26	608.5	22.7	454	2	AAR84551	Partial h
27	608.5	22.7	454	3	AAY94009	Amino aci
28	601.5	22.5	437	7	ADD45362	Rat Prote
29	572.5	21.4	437	7	ADD69273	Human fib
30	572.5	21.4	453	7	ADD69274	Human fib
31	544.5	20.3	831	8	ADL92161	Fibrinoge
32	544.5	20.3	847	3	AY82891	Alphab su
33	544.5	20.3	866	7	ADD48695	Human Pro
34	539.5	20.2	251	3	AY82888	AlphaECX
35	537.5	20.1	236	3	AY82889	AlphaECX
36	531.5	19.9	493	3	AAB41712	Human ORF
37	531.5	19.9	493	6	ABR55568	Amino aci
38	530.5	19.8	493	2	AY05396	Human TIE
39	530.5	19.8	493	3	AY54999	Human sca
40	530.5	19.8	493	3	AY70745	PSEQ-3 pr
41	530.5	19.8	493	3	AAB47997	Angiopoie
42	530.5	19.8	493	4	AAB31179	Amino aci
43	530.5	19.8	493	4	AAL12305	Human PRO
44	530.5	19.8	493	4	AAB51329	Human 410
45	530.5	19.8	493	5	AAO14786	Human mat

QY 121 RPIRNSVDELNNVAVSOTSSSFQYMYLLKDLWQKQKQVNDNENNVNVEYSSELEKHQ 180
 DB 121 RPIRNSVDELNNVAVSOTSSSFQYMYLLKDLWQKQKQVNDNENNVNVEYSSELEKHQ 180
 QY 181 LYIDETVNSNIATNLRLVLSILENLSRKIQKLESVSAQMEYCRTPCTVSCNIPVVSQKE 240
 DB 181 LYIDETVNSNIPTNLRLVLSILENLSRKIQKLESVSAQMEYCRTPCTVSCNIPVVSQKE 240
 QY 241 CEELIRKGETSEMYLIQPDSSVKPYRVVCDMNTENGWTVIQRQDGSVDFGRKWDPYK 300
 DB 241 CEELIRKGETSEMYLIQPDSSVKPYRVVCDMNTENGWTVIQRQDGSVDFGRKWDPYK 300
 QY 301 QFGNVATNTDGNKYCGLPGEYWLGNDKISQLRMGPTTELLIEMEDWKGDVKVAHYGGFT 360
 DB 301 QFGNVATNTDGNKYCGLPGEYWLGNDKISQLRMGPTTELLIEMEDWKGDVKVAHYGGFT 360
 QY 361 VQNEANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGWLTSDF 420
 DB 361 VQNEANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGWLTSDF 420
 QY 421 RKQCKEDGGWYNRCHAANPNRGYYWGQYTWDMAKHGTDGCVVMWKNKGSWYSMRKM 480
 DB 421 RKQCKEDGGWYNRCHAANPNRGYYWGQYTWDMAKHGTDGCVVMWKNKGSWYSMRKM 480
 QY 481 SMKIRPFFPQQ 491
 DB 481 SMKIRPFFPQQ 491

RESULT 6
 AAM79475
 ID AAM79475 standard; protein; 495 AA.
 AC AAM79475;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 3121.
 DE
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US0004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR N-PSDB; AAK52608.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 in diagnosis and gene therapy.

XX
 PS
 XX
 CC Claim 20; Page 257-258; 6221pp; English.
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. the polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 495 AA;
 Query Match 99.1%; Score 2651; DB 4; Length 495;
 Best Local Similarity 99.4%; Pred. No. 1.1e-212;
 Matches 486; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 RMVVSFHLKTKMKHLLLLLCVFLVKSQGVNDNEGFFSARGHRPLDKKREAPSLRPA 62
 DB 7 RPVVSFHLKTKMKHLLLLLCVFLVKSQGVNDNEGFFSARGHRPLDKKREAPSLRPA 66
 QY 63 PPIISGGYRARPAAATQKKVERKAPDAGCLHADPLGVLCPTGCOLQEQERP 122
 DB 67 PPIISGGYRARPAAATQKKVERKAPDAGCLHADPLGVLCPTGCOLQEQERP 126
 QY 123 IRNSVDELNNVAVSOTSSSFQYMYLLKDLWQKQKQVNDNENNVNVEYSSELEKHLY 182
 DB 127 IRNSVDELNNVAVSOTSSSFQYMYLLKDLWQKQKQVNDNENNVNVEYSSELEKHLY 186
 QY 183 IDETVNSNIATNLRLVLSILENLSRKIQKLESVSAQMEYCRTPCTVSCNIPVVSQKE 242
 DB 187 IDETVNSNIPTNLRLVLSILENLSRKIQKLESVSAQMEYCRTPCTVSCNIPVVSQKE 246
 QY 243 EIIRKGETSEMYLIQPDSSVKPYRVVCDMNTENGWTVIQRQDGSVDFGRKWDPYKQ 302
 DB 247 EIIRKGETSEMYLIQPDSSVKPYRVVCDMNTENGWTVIQRQDGSVDFGRKWDPYKQ 306
 QY 303 FGNVATNTDGNKYCGLPGEYWLGNDKISQLRMGPTTELLIEMEDWKGDVKVAHYGGFTVQ 362
 DB 307 FGNVATNTDGNKYCGLPGEYWLGNDKISQLRMGPTTELLIEMEDWKGDVKVAHYGGFTVQ 366
 QY 363 NEANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGWLTSDFPK 422
 DB 367 NEANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGWLTSDFPK 426
 QY 423 QCKSKEDGGWYNRCHAANPNRGYYWGQYTWDMAKHGTDGCVVMWKNKGSWYSMRKM 482
 DB 427 QCKSKEDGGWYNRCHAANPNRGYYWGQYTWDMAKHGTDGCVVMWKNKGSWYSMRKM 486
 QY 483 KIRPFFPQQ 491
 DB 487 KIRPFFPQQ 495

RESULT 7
 AAM79476
 ID AAM79476 standard; protein; 495 AA.
 XX
 AC AAM79476;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 3122.
 DE
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;

CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 495 AA;

Query Match 99.1%; Score 2651; DB 4; Length 495;
Best Local Similarity 99.4%; Pred. No. 1.1e-212;
Matches 486; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 RMVSWFHKLTWKHLLLLCVLVKSGVNDNEGFFSARGHRPLDKKREAPSRLPA 62
Db 7 RPYWSFHKLTWKHLLLLCVLVKSGVNDNEGFFSARGHRPLDKKREAPSRLPA 66
QY 63 PPISSGGYRPAKAAATOKKVERKAPDAGGCLHADPDGLVLCPTGCOLQOALLQOERP 122
Db 67 PPISSGGYRPAKAAATOKKVERKAPDAGGCLHADPDGLVLCPTGCOLQOALLQOERP 126
QY 123 IRNSVDELNNNEAVSQTSSSFQYMYLLKDLWKQKQVKNENNVNVEYSSELEKHLY 182
Db 127 IRNSVDELNNNEAVSQTSSSFQYMYLLKDLWKQKQVKNENNVNVEYSSELEKHLY 186
QY 183 IDTVNSNATNLRVLRSLILENRSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSKECE 242
Db 187 IDTVNSNATNLRVLRSLILENRSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSKECE 246
QY 243 EIIRKGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQRDQGSVDFGRKWDYKQG 302
Db 247 EIIRKGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQRDQGSVDFGRKWDYKQG 306
QY 303 FGNVATNTDGNKCYGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQ 362
Db 307 FGNVATNTDGNKCYGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQ 366
QY 363 NEANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGLTSDPRK 422
Db 367 NEANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGLTSDPRK 426
QY 423 QSKEDGGGWWYNRCHAANPNRGYYWGGQYTDMAKHGTTDGGVVMNWKGSWYSMEKMSM 482
Db 427 QSKEDGGGWWYNRCHAANPNRGYYWGGQYTDMAKHGTTDGGVVMNWKGSWYSMEKMSM 486
QY 483 IRPFPFQQ 491
Db 487 IRPFPFQQ 495

RESULT 9
ADE76868
ID ADE76868 standard; protein; 488 AA.
AC ADE76868;
XX
DT 29-JAN-2004 (first entry)
DE Human protein expressed in a liver disorder #11.
XX
KW human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
KW high-throughput screening; differential gene expression; gene therapy.
XX
OS Homo sapiens.

XX
XX US2003108871-A1.
XX
PD 12-JUN-2003.
XX
PF 30-JUL-2001; 2001US-00919039.
XX
PR 28-JUL-2000; 2000US-0222113P.
XX
PA (KASE/) KASER M R.

PI Kaser MR;
XX WPI; 2004-031227/03.
DR N-PSDB; ADE76867.
XX
PT Composition comprising several cDNAs that are differentially expressed in
PT treated human C3A liver cell cultures, useful for treating liver
PT disorders.
XX
XX Claim 1; SEQ ID NO 33; 41pp; English.

PS The invention relates to a composition comprising several cDNAs that are
XX differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound, therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a
CC liver disorder.

XX Sequence 488 AA;

Query Match 99.0%; Score 2648; DB 8; Length 488;
Best Local Similarity 99.4%; Pred. No. 2e-212;
Matches 485; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 MYSWSPHKLTWKHLLLLCVLVKSGVNDNEGFFSARGHRPLDKKREAPSRLPAP 63
Db 1 MYSWSPHKLTWKHLLLLCVLVKSGVNDNEGFFSARGHRPLDKKREAPSRLPAP 60
QY 64 PPISSGGYRPAKAAATOKKVERKAPDAGGCLHADPDGLVLCPTGCOLQOALLQOERP 123
Db 61 PPISSGGYRPAKAAATOKKVERKAPDAGGCLHADPDGLVLCPTGCOLQOALLQOERP 120
QY 124 IRNSVDELNNNEAVSQTSSSFQYMYLLKDLWKQKQVKNENNVNVEYSSELEKHLY 183
Db 121 IRNSVDELNNNEAVSQTSSSFQYMYLLKDLWKQKQVKNENNVNVEYSSELEKHLY 180
QY 184 IDTVNSNATNLRVLRSLILENRSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSKECE 243
Db 181 IDTVNSNATNLRVLRSLILENRSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSKECE 240
QY 244 EIIRKGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQRDQGSVDFGRKWDYKQG 303
Db 241 EIIRKGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQRDQGSVDFGRKWDYKQG 300
QY 304 GNVATNTDGNKCYGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQ 363
Db 301 GNVATNTDGNKCYGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQ 360
QY 364 EANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGLTSDPRK 423
Db 361 EANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGLTSDPRK 420
QY 424 QSKEDGGGWWYNRCHAANPNRGYYWGGQYTDMAKHGTTDGGVVMNWKGSWYSMEKMSM 483
Db 421 QSKEDGGGWWYNRCHAANPNRGYYWGGQYTDMAKHGTTDGGVVMNWKGSWYSMEKMSM 480
QY 484 IRPFPFQQ 491
Db 481 IRPFPFQQ 488

```

RESULT 10
AAM78491
ID AAM78491 standard; protein; 539 AA.
XX
AC AAM78491;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1153.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK51624.
XX
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 3388-3389; 6221pp; English.
XX
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 539 AA;
Query Match 98.5%; Score 2637; DB 4; Length 539;
Best Local Similarity 90.9%; Pred. No. 1.9e-211;
Matches 490; Conservative 0; Mismatches 1; Indels 48; Gaps 1;
XX
QY 1 MKRMVSVFHLKTMKTHLLLLLCVFLVKSQGVNDNEE-----38
DB 1 MKRMVSVFHLKTMKTHLLLLLCVFLVKSQGVNDNEEQYRITKTRSENLTNYKIIE 60
QY 39 -----CFESARGHRPLDKREAPSLRAPPPIISGGGYR 72

```


DE	Human protein SEQ ID NO 1155.		
XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
KW	nervous system disorder; arthritis; inflammation.		
OS	Homo sapiens.		
XX			
XX	WO200157190-A2.		
XX	09-AUG-2001.		
XX			
XX	05-FEB-2001; 2001WO-US004098.		
XX	03-FEB-2000; 2000US-00496914.		
PR	27-APR-2000; 2000US-00508075.		
PR	20-JUN-2000; 2000US-00598075.		
PR	19-JUL-2000; 2000US-00620325.		
PR	01-SEP-2000; 2000US-00654936.		
PR	15-SEP-2000; 2000US-00663561.		
PR	20-OCT-2000; 2000US-00693325.		
PR	30-NOV-2000; 2000US-00728422.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;		
PI	Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;		
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;		
XX			
DR	WPI; 2001-476283/51.		
DR	N-PSDB; AAK51626.		
XX			
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful		
PT	in diagnosis and gene therapy.		
XX			
PS	Claim 20; Page 3390-3391; 6221pp; English.		
XX			
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the		
CC	encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	e.g. stem cell growth factor activity, haematopoiesis regulating		
CC	activity, tissue growth factor activity, immunomodulatory activity and		
CC	activin/inhibin activity and may be useful in the diagnosis and/or		
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and		
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111		
CC	(AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the		
CC	sequence listing were missing at the time of publication		
XX			
SQ	Sequence 453 AA;		
Query Match	31.0%; Score 2436; DB 4; Length 453;		
Best Local Similarity	92.1%; Pred. No. 9.9e-195;		
Matches 452; Conservative	0; Mismatches 1; Indels 38; Gaps 1;		
QY	1 MKRMVSWSHKLTMTKHLILLLCVFLVKSQGVNDNEEGFFSARGHRPLDKKREAPSLR 60		
DB	1 MKRMVSWSHKLTMTKHLILLLCVFLVKSQGVNDNEEGFFSARGHRPLDKKREAPSLR 60		
QY	61 PAPPTISGGGYRARPAAKAAATKKVERKAPDAGGCLHADPDGLVLCPTGQQLQOEALLQOE 120		
DB	61 PAPPTISGGGYRARPAAKAAATKKVERKAPDAGGCLHADPDGLVLCPTGQQLQOEALLQOE 120		
QY	121 RPTNSVDELNNVNAVVSOTSSSFQMYLLDLWKQKQKQKQKQKQKQKQKQKQKQKQKQ 180		
DB	121 RPTNSVDELNNVNAVVSOTSSSFQMYLLDLWKQKQKQKQKQKQKQKQKQKQKQKQKQ 180		
QY	181 LYIDETVNSNIATNLRLVLSILENLRSKIQLKESDVSQAQMEYCRTPCTVSCNIPVVSQKE 240		
DB	181 LYIDETVNSNIPTNLRLVLSILENLRSKIQLKESDVSQAQMEYCRTPCTVSCNIPVVSQK- 239		
QY	241 CEEIIRKGETSEMYLIQPDSSVKPYRYVCDMNTENGWTVIQNRQDGSVDGFRKWDPKYK 300		
DB	240 -----GWTVIQNRQDGSVDGFRKWDPKYK 262		
QY	301 QGFGNVATNDGKNYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDMKDKYKAHYGGFT 360		
DB	263 QGFGNVATNDGKNYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDMKDKYKAHYGGFT 322		
QY	361 VONEANKYQISVKNKYRGTTAGNALMDGASQLMGENTMTIHNMGFFSTYDRDNDGWLTSDP 420		
DB	323 VONEANKYQISVKNKYRGTTAGNALMDGASQLMGENTMTIHNMGFFSTYDRDNDGWLTSDP 382		
QY	421 RXQCSKEDGGGWWYNRCHAAANPNGRYWGQVYTDMAKHGTDDGVVMMNWKGSWYSMRKM 480		
DB	383 RXQCSKEDGGGWWYNRCHAAANPNGRYWGQVYTDMAKHGTDDGVVMMNWKGSWYSMRKM 442		
QY	481 SMKIRPFPFQQ 491		
DB	443 SMKIRPFPFQQ 453		
RESULT 14			
ADI82107			
ID	ADI82107 standard; protein; 479 AA.		
XX			
AC	ADI82107;		
XX			
DT	22-APR-2004 (first entry)		
XX			
DE	Fibrinogen-beta-chain protein.		
XX			
KW	kidney toxicity; T-kininogen; inter-alpha-inhibitor H4P heavy chain;		
KW	alpha-1-macroglobulin; apolipoprotein E; clusterin; complement C3;		
KW	complement C4; fibrinogen-alpha-chain; fibrinogen-beta-chain;		
KW	plasma retinol binding protein; renal tubular necrosis; drug development.		
XX			
OS	Unidentified.		
XX			
PN	WO2004005934-A2.		
XX			
PD	15-JAN-2004.		
XX			
PF	04-JUL-2003; 2003WO-GB002893.		
XX			
PR	04-JUL-2002; 2002GB-00015575.		
PR	09-SEP-2002; 2002GB-00020879.		
XX			
PA	(OXFO-) OXFORD GLYCOSCIENCES UK LTD.		
XX			
PI	Ranasinghe-Bandara L;		
XX			
DR	WPI; 2004-099412/10.		
XX			
PT	Screening or diagnosing kidney toxicity, useful in monitoring		
PT	effectiveness of treatment for kidney toxicity or for screening and		
PT	developing of drugs, by detecting and quantifying T-kininogen and inter-		
PT	alpha-inhibitor H4P heavy chain.		
XX			
PS	Disclosure; SEQ ID NO 10; 41pp; English.		
XX			
CC	The invention comprises a method for screening or diagnosing kidney		
CC	toxicity. The method involves the use of the following proteins: 1-		
CC	kininogen,inter-alpha-inhibitor H4P heavy chain, alpha-1-macroglobulin,		
CC	apolipoprotein E, clusterin, complement C3, complement C4, fibrinogen-		
CC	alpha-chain, fibrinogen-beta-chain, and plasma retinol binding protein.		
CC	The method of the invention is useful in screening or diagnosing kidney		
CC	toxicity (e.g. renal tubular necrosis). The method is also useful for		
CC	monitoring the effectiveness of treatment for kidney toxicity and for		
CC	screening and developing drugs. The present amino acid sequence		
CC	represents a protein that was used in the exemplification of the		
CC	invention.		

Time : 169 secs

Pt cancer and non-cancerous diseases of the lung.
XX

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:20 ; Search time 2161.38 Seconds
(without alignments)
11849.778 Million cell updates/sec

Title: US-10-017-724-5_COPY_4000_8878

Perfect score: 4879

Sequence: 1 tcagaataagtggtgatgag.....gattgcattgagccaagatc 4879

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	4879	100.0	8878	2	AAT03852 Human fib
2	4879	100.0	8878	6	AAD45341 Human ant
3	4879	100.0	8878	9	AAD51117 Human FBG
4	4879	100.0	8878	11	ADP65553 Human fib
5	4879	100.0	8878	11	ADP65677 Human fib
6	4268.2	87.5	8674	8	ABX63814 Human cDN
7	2081.2	42.7	4043	6	ABX91986 Lung spec
8	971.8	19.9	4967	5	AAS60957 Human can
9	971.8	19.9	4967	5	ADL63650 Human ova
10	659.6	13.5	2553	4	AAS22608 Human pol
11	659.6	13.5	2553	4	AAS22609 Human pol
12	659.6	13.5	2553	4	AAS22610 Human pol
13	658.6	13.5	1862	4	AAS51626 Human pol
14	658.6	13.5	1976	4	AAS51625 Human pol
15	658.6	13.5	2120	4	AAS51624 Human pol
16	656.4	13.5	1918	11	ADP65102 Human fib
17	656.4	13.5	1918	11	ADP65031 Human fib
18	653.4	13.4	1883	11	ADP65676 Human fib
19	653.4	13.4	1883	11	ADP65552 Human fib
20	594	12.2	2553	4	AAS22608 Human pol
21	594	12.2	2553	4	AAS22609 Human pol

C	22	594	12.2	2553	4	AAS22610 Human pol
	23	454.6	9.3	455	3	AAC71409 Single nu
	24	454.6	9.3	455	3	AAC71418 Single nu
	25	454.6	9.3	455	3	AAC71443 Single nu
	26	447.6	9.2	448	3	AAC71412 Single nu
	27	447.6	9.2	448	3	AAC71421 Single nu
	28	430.6	8.8	431	3	AAC71427 Single nu
	29	397.2	8.1	414	9	ACH20725 Human adu
	30	390.4	8.0	404	9	ACH20470 Human adu
	31	387.8	7.9	415	9	ACH20471 Human adu
	32	387.8	7.9	1629	12	ADE76867 Human cDN
	33	369.4	7.6	399	9	ACH21083 Human adu
	34	368.4	7.6	382	9	ACH20960 Human adu
C	35	362.2	7.4	571	6	ABN96686 Gene #318
	36	338.2	6.9	375	6	ABN96703 Gene #320
	37	335.6	6.9	385	9	ACH21036 Human adu
	38	324.4	6.6	388	9	ACH21014 Human adu
	39	316.6	6.5	317	3	AAC71403 Single nu
	40	288.6	5.9	413	9	ACH20958 Human adu
	41	284.6	5.8	285	3	AAC71433 Single nu
	42	284.6	5.8	285	3	AAC71424 Single nu
	43	277.2	5.7	306	9	ACH20753 Human adu
	44	275.6	5.6	412	9	ACH20888 Human adu
C	45	264.4	5.4	357	6	ABN94253 Gene #751

ALIGNMENTS

RESULT 1

AAT03852

ID AAT03852 standard; DNA; 8878 BP.

XX AC AAT03852;

XX DT 26-OCT-1996 (first entry)

XX DE Human fibrinogen B-beta chain coding sequence.

XX KW Human fibrinogen B-beta chain; transgenic animal milk; treatment;

XX KW sheep beta-lactoglobulin signal peptide fusion protein;

XX KW surgical adhesive; ds.

XX OS Homo sapiens.

XX FH Key

FT misc_RNA

FT CDS

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

Location/Qualifiers
1. .469
/*tag= a
470. .8102
/*tag= b
584. .3257
/*tag= c
/*tag= "I"
3258. .3449
/*tag= d
/*tag= "2"
3450. .3938
/*tag= e
/*tag= "2"
3939. .4122
/*tag= f
/*tag= "3"
4123. .5042
/*tag= g
/*tag= "3"
5043. .5270
/*tag= h
/*tag= "4"
5271. .5830

```

FT      /*tag= i
FT      /note= "4"
FT      exon      5831..5944
FT      /*tag= j
FT      /note= "5"
FT      intron    5945..6632
FT      /*tag= k
FT      /note= "5"
FT      exon      6633..6758
FT      /*tag= l
FT      /note= "6"
FT      intron    6759..6966
FT      /*tag= m
FT      /note= "6"
FT      exon      6967..7252
FT      /*tag= n
FT      /note= "7"
FT      intron    7253..7870
FT      /*tag= o
FT      /note= "7"
FT      exon      7871..8102
FT      /*tag= p
FT      /note= "8"
FT      3'UTR     8103..8537
FT      /*tag= q
FT      misc_RNA 8538..8878
FT      /*tag= r
XX
PN WO9523868-A1.
XX
XX      08-SEP-1995.
XX
XX      01-MAR-1995; 95WO-US002648.
XX
XX      03-MAR-1994; 94US-00206176.
XX
XX      (ZYMO ) ZYMOGENETICS INC.
XX      (PHAR-) PHARM PROTEINS LTD.
XX
XX      Garner I, Dalrymple MA, Prunkard DE, Foster DC;
XX      WPI; 1995-320582/41.
XX      P-PSDB; AAR82243.
XX
XX      Production of fibrinogen in transgenic mammals - by introducing DNA
XX      segments into the germ line of a non-human mammal and collecting milk
XX      from female progeny.
XX
XX      Claim 9; Page 47-55; 99pp; English.
XX
XX      This sequence encodes the human fibrinogen B-beta chain. Preferably
XX      nucleotides 470-8100, and more preferably nucleotides 512-8100, are
XX      operably linked to the signal peptide, preferably of the sheep beta-
XX      lactoglobulin gene (see AAR03855) and, together with the fibrinogen A-
XX      alpha chain sequence (see AAR03853) and the gamma chain sequence (see
XX      AAR03854), is introduced into a fertilised egg or the germ line of a non-
XX      human animal, preferably a sheep, pig, goat or cattle at a molar ratio of
XX      0.5-1.0-0.5-1.0-0.5-1.0, respectively, for fusion protein gene expression
XX      in transgenic animal milk. The recombinantly produced active fibrinogen
XX      is useful in human and veterinary medicine, e.g. in the formulation of
XX      surgical adhesives, which also consist of Factor-XIII, and as a coating
XX      surface for polymeric articles, such as synthetic vascular grafts
XX
XX      Sequence 8878 BP; 2912 A; 1477 C; 1664 G; 2825 T; 0 U; 0 Other;
XX
XX      Query Match      100.0%; Score 4879; DB 2; Length 8878;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      1 TCAGAAATAGTGTGATGAGTTAAATACAAATGTGGAAGCTGTTCCAGACCTCTCTT 60
XX      |||||||
XX      4000 TCAGAAATAGTGTGATGAGTTAAATACAAATGTGGAAGCTGTTCCAGACCTCTCTT 4059
XX      |||||||

```

Db 5140 TGTGCTTCGTTCAATCTCGGAAACCTGAGAAAGCAAAATACAAAAGTTAGAACTCTGATG 5199
Qy 1201 CTCAGCTCAAAATGGAATATGTCGCAACCCCATGCACTGCTGAGTTGCAATATTCCTGTGGT 1260
Db 5200 CTCAGCTCAAAATGGAATATGTCGCAACCCCATGCACTGCTGAGTTGCAATATTCCTGTGGT 5259
Qy 1261 GTCTGGCAAGGTAACTGATTCATATAACATATTTTAGAGAGTTCCAGAAAGACTCACAC 1320
Db 5260 GTCTGGCAAGGTAACTGATTCATATAACATATTTTAGAGAGTTCCAGAAAGACTCACAC 5319
Qy 1321 ACCAAAATAAGAGAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1380
Db 5320 ACCAAAATAAGAGAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 5379
Qy 1381 ATAATGACATTTACAGTACATCATATAAAATATCCTTTAGCCAGTTGTTTGGACTGGCCT 1440
Db 5380 ATAATGACATTTACAGTACATCATATAAAATATCCTTTAGCCAGTTGTTTGGACTGGCCT 5439
Qy 1441 GGTGCAATTTGCTGGTTTGTATGAGCAGGATGGGCAACAGTGTAGTCCAGGGTGGCTGAT 1500
Db 5440 GGTGCAATTTGCTGGTTTGTATGAGCAGGATGGGCAACAGTGTAGTCCAGGGTGGCTGAT 5499
Qy 1501 GTGTGCATCTCGGTACTGGCTTGAACAGATGGCAGAACCAACAGATAGATGTAGAAAGTTTC 1560
Db 5500 GTGTGCATCTCGGTACTGGCTTGAACAGATGGCAGAACCAACAGATAGATGTAGAAAGTTTC 5559
Qy 1561 TCCATTTTGTGTGTTCTGGAGCTCATGGATATCCAGAGCAACAAAGGTGGAGAGGC 1620
Db 5560 TCCATTTTGTGTGTTCTGGAGCTCATGGATATCCAGAGCAACAAAGGTGGAGAGGC 5619
Qy 1621 TTTGTTTCATCTTACAGATTAACAGTCTCAAACTGGGTGGACTTACTTAAGATAAA 1680
Db 5620 TTTGTTTCATCTTACAGATTAACAGTCTCAAACTGGGTGGACTTACTTAAGATAAA 5679
Qy 1681 ATGAAATCTCAATATTTGTTATTAATTTTCAAAGTCTATAATAACACACTCCTTAGTA 1740
Db 5680 ATGAAATCTCAATATTTGTTATTAATTTTCAAAGTCTATAATAACACACTCCTTAGTA 5739
Qy 1741 ACTTATGTAATGTTATTTTAAAGAAATTTGGTGACTAAATAACAAAGTAAATATGTCTATAAAC 1800
Db 5740 ACTTATGTAATGTTATTTTAAAGAAATTTGGTGACTAAATAACAAAGTAAATATGTCTATAAAC 5799
Qy 1801 CCTGAAACATAATGTTGTTCTTACATTTGCAGATGTGAGGAAATATCAGGAAAGAGGT 1860
Db 5800 CCTGAAACATAATGTTGTTCTTACATTTGCAGATGTGAGGAAATATCAGGAAAGAGGT 5859
Qy 1861 GAAACATCTGAAATGATCTCATTCAACCTGCAGTCTCTGCAACCGTATAGATATAC 1920
Db 5860 GAAACATCTGAAATGATCTCATTCAACCTGCAGTCTCTGCAACCGTATAGATATAC 5919
Qy 1921 TGTGACATGAATACAGAAATGAGGTAAAGCTTTTCAGCAGTTGTTGCACTGTTGATCTGT 1980
Db 5920 TGTGACATGAATACAGAAATGAGGTAAAGCTTTTCAGCAGTTGTTGCACTGTTGATCTGT 5979
Qy 1981 AATTTTGGATACCGTAAATGCGCAGGAAAACAGGCCAGGTGTGGCTCATACCTGT 2040
Db 5980 AATTTTGGATACCGTAAATGCGCAGGAAAACAGGCCAGGTGTGGCTCATACCTGT 6039
Qy 2041 AATTTCCAGCACCTTTGGAGGCCAAAGTGGCTCATAGCTTTGAGCTTAGGAGTTTGAACCT 2100
Db 6040 AATTTCCAGCACCTTTGGAGGCCAAAGTGGCTCATAGCTTTGAGCTTAGGAGTTTGAACCT 6099
Qy 2101 AGCTGGGCAACATAATGAGACCTTAACCTCAACAAAAAATAACAAAAAATAACAAAAAATA 2160
Db 6100 AGCTGGGCAACATAATGAGACCTTAACCTCAACAAAAAATAACAAAAAATAACAAAAAATA 6159
Qy 2161 AAAAAATACGCTGTGTTGTTAGTATGTGCTGTAGTCCAGCTATCCAGGAGGCTGAGA 2220
Db 6160 AAAAAATACGCTGTGTTGTTAGTATGTGCTGTAGTCCAGCTATCCAGGAGGCTGAGA 6219
Qy 2221 TGGGAGATCACCTGAGGCCAACCTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGG 2280

Db 6220 TGGGAGATCACCTGAGGCCCAACACCTGGAGTCTTGTGATCATGCTACTGAACCTGTAGCCTGG 6279
Qy 2281 GCAACAGAGGTACTGAGATCCTGCTCTCAAAAAAATAAATTTAAAAAGCCAGGAAAC 2340
Db 6280 GCAACAGAGGTACTGAGATCCTGCTCTCAAAAAAATAAATTTAAAAAGCCAGGAAAC 6339
Qy 2341 AAGACTTAGCTCTAACATCTAACATAGCTGACAAAGGAGTAAATTTGATGTGGAATTCAC 2400
Db 6340 AAGACTTAGCTCTAACATCTAACATAGCTGACAAAGGAGTAAATTTGATGTGGAATTCAC 6399
Qy 2401 CTGATTTTAAAAAGTTATATAATATCTATAATCTCAAAATTTGGGTAAGATAAGACATT 2460
Db 6400 CTGATTTTAAAAAGTTATATAATATCTATAATCTCAAAATTTGGGTAAGATAAGACATT 6459
Qy 2461 CGAGTTTCCAAAGATTTTACAGTTTACCTCTCATATTTATTTCCCTTATTTGTCATTT 2520
Db 6460 CGAGTTTCCAAAGATTTTACAGTTTACCTCTCATATTTATTTCCCTTATTTGTCATTT 6519
Qy 2521 TAGAGCAACAAATATATATACTAAATGGAATGGAACAGGGATTCAGATATTTTCCAAAGT 2580
Db 6520 TAGAGCAACAAATATATATACTAAATGGAATGGAACAGGGATTCAGATATTTTCCAAAGT 6579
Qy 2581 GACATTTTGTGTTGTTAATATATGCTCTTTTTGTTTCTGTCTCAACAAAGGATGGAC 2640
Db 6580 GACATTTTGTGTTGTTAATATATGCTCTTTTTGTTTCTGTCTCAACAAAGGATGGAC 6639
Qy 2641 AGTGATTCAGAACCGTCAAGCGTAGTTGACTTTTCGCAGGAATGGGATCCATATAA 2700
Db 6640 AGTGATTCAGAACCGTCAAGCGTAGTTGACTTTTCGCAGGAATGGGATCCATATAA 6699
Qy 2701 ACAGGGATTTGGAATGTTGCAACCAACACACAGATGGGAAGATTTACTGTGGCTACCAGG 2760
Db 6700 ACAGGGATTTGGAATGTTGCAACCAACACACAGATGGGAAGATTTACTGTGGCTACCAGG 6759
Qy 2761 TAAACAAACAGGCATCAAAATAAAATCAATCTTATTTGAAATGGGATTTTTTTTAAATAAA 2820
Db 6760 TAAACAAACAGGCATCAAAATAAAATCAATCTTATTTGAAATGGGATTTTTTTTAAATAAA 6819
Qy 2821 AAAACATTCATTTGTTGGAGCCTGTTTATAGGCAAGTAAAGGAGTTTCCGTGACAAAATCT 2880
Db 6820 AAAACATTCATTTGTTGGAGCCTGTTTATAGGCAAGTAAAGGAGTTTCCGTGACAAAATCT 6879
Qy 2881 GGAAGCTAAAGATAAGGGAAGAAAGCCAGTCTTTTATAGTTTCCCAAAATTTTATTTTGGTG 2940
Db 6880 GGAAGCTAAAGATAAGGGAAGAAAGCCAGTCTTTTATAGTTTCCCAAAATTTTATTTTGGTG 6939
Qy 2941 AGAGATTTTATTTTGTGTTTCTTTTATAGGTAATAATGCGCTTGGAAATGATAAAATAGCC 3000
Db 6940 AGAGATTTTATTTTGTGTTTCTTTTATAGGTAATAATGCGCTTGGAAATGATAAAATAGCC 6999
Qy 3001 AGCTTACAGGATGGGACCCACAGACTTTTTCATAGAAATGGAGACTGGAAGGAGACA 3060
Db 7000 AGCTTACAGGATGGGACCCACAGACTTTTTCATAGAAATGGAGACTGGAAGGAGACA 7059
Qy 3061 AAGTAAAGGCTCACTATGAGGATTCACGTGTACAGAAATGAAGCCAAACAAATACCAGATCT 3120
Db 7060 AAGTAAAGGCTCACTATGAGGATTCACGTGTACAGAAATGAAGCCAAACAAATACCAGATCT 7119
Qy 3121 CAGTGAACAAATAACAGAGAAACAGCGGTAATGCCCTCATGATGGAGCATCTCAGCTGA 3180
Db 7120 CAGTGAACAAATAACAGAGAAACAGCGGTAATGCCCTCATGATGGAGCATCTCAGCTGA 7179
Qy 3181 TGGGAGAAAAACAGGACCATGACCATTCACAAAGGATGTTCTTCAGCAGTATGACAGAG 3240
Db 7180 TGGGAGAAAAACAGGACCATGACCATTCACAAAGGATGTTCTTCAGCAGTATGACAGAG 7239
Qy 3241 ACAATGACCGCTGGTATGTGTGCTCTTGTCTCTGCTTTAAAAATCACTAATATC 3300
Db 7240 ACAATGACCGCTGGTATGTGTGCTCTTGTCTCTGCTTTAAAAATCACTAATATC 7299
Qy 3301 ATTACTCAGAATCATTTAAACAATAATTTTAAATAGTACCACTTCTCGGGCACTTACTGTCA 3360
Db 7300 ATTACTCAGAATCATTTAAACAATAATTTTAAATAGTACCACTTCTCGGGCACTTACTGTCA 7359

CC containing a flour, extract or malt obtained from mature monocot seeds
CC and one or more heterologous seed-produced anti-microbial proteins in
CC substantially unpurified form. The invention is useful as a feed for
CC production animals such as poultry and hoofed farm animals. The present
CC sequence is human anti-microbial protein encoding DNA. This sequence is
CC used in the invention
XX
SQ Sequence 8878 BP; 2912 A; 1478 C; 1663 G; 2825 T; 0 U; 0 Other;

Query Match 100.0%; Score 4879; DB 6; Length 8878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAAATAGTGTGATGATGATTAATAACAAGTGAAGCTGTTCCAGAGCTCTCTT 60
DB 4000 TCAGAAATAGTGTGATGATGATTAATAACAAGTGAAGCTGTTCCAGAGCTCTCTT 4059

QY 61 CTTCTTTTCAGTACATGATTTTGTCTGAAAGACCTGTGGCAAGAGGCGAAGCAAGTAA 120
DB 4060 CTTCTTTTCAGTACATGATTTTGTCTGAAAGACCTGTGGCAAGAGGCGAAGCAAGTAA 4119

QY 121 AAGGTAGATATCTTGTGCTTCCATTCGATTTTCAGCTTATAAATGGAAACCGTTAGAC 180
DB 4120 AAGGTAGATATCTTGTGCTTCCATTCGATTTTCAGCTTATAAATGGAAACCGTTAGAC 4179

QY 181 TGCCACGAGAAATGCAATGTTGTGGAAGATTAACATTTCTGGTGTAGTCAATGCAATCA 240
DB 4180 TGCCACGAGAAATGCAATGTTGTGGAAGATTAACATTTCTGGTGTAGTCAATGCAATCA 4239

QY 241 TACGCTTTTGGGCACCTTCCCTGCAACTTGGCAGATAGCACTATTCAAGCTCTTATTC 300
DB 4240 TACGCTTTTGGGCACCTTCCCTGCAACTTGGCAGATAGCACTATTCAAGCTCTTATTC 4299

QY 301 CAGTCTGACATCAGCAAGTGTGATTTCTATGAAATTTCTACTAGCTCTCTTATTTTA 360
DB 4300 CAGTCTGACATCAGCAAGTGTGATTTCTATGAAATTTCTACTAGCTCTCTTATTTTA 4359

QY 361 AGTATACAGAAATCTGTGACCTCAGAGATTAATTTACAGAGTGGAAAAACCCCTAG 420
DB 4360 AGTATACAGAAATCTGTGACCTCAGAGATTAATTTACAGAGTGGAAAAACCCCTAG 4419

QY 421 CATTTATAGTTTAAATTTGAGGTTTGAATGAGAGAGTATCCATATATATTAATTTCA 480
DB 4420 CATTTATAGTTTAAATTTGAGGTTTGAATGAGAGAGTATCCATATATATTAATTTCA 4479

QY 481 GTGTTGTGATTAATGACACTAACTGTGAAATCTTGAGGTCAAGATTTGAGTGTGTTG 540
DB 4480 GTGTTGTGATTAATGACACTAACTGTGAAATCTTGAGGTCAAGATTTGAGTGTGTTG 4539

QY 541 ACTTGGTGTGAGGAAACAGCTAGTGGCTGGCTGGCAGGCACTCTCAGTGAAGTAC 600
DB 4540 ACTTGGTGTGAGGAAACAGCTAGTGGCTGGCTGGCAGGCACTCTCAGTGAAGTAC 4599

QY 601 TACCACAGTTGGAAATTTTTCAGAGAAATCAAAGAAATCATGACATCTTATAAATTTCA 660
DB 4600 TACCACAGTTGGAAATTTTTCAGAGAAATCAAAGAAATCATGACATCTTATAAATTTCA 4659

QY 661 AGGTTCTGTCTACTTATGTAATGGAATAAATAAATCAAGCATATPCCACTCTGAAGAT 720
DB 4660 AGGTTCTGTCTACTTATGTAATGGAATAAATAAATCAAGCATATPCCACTCTGAAGAT 4719

QY 721 TGAACCTTCTCAGATGGAAGACCCCAATCTGCTTCTCTCTCTCTCTCTCTCTCTCT 780
DB 4720 TGAACCTTCTCAGATGGAAGACCCCAATCTGCTTCTCTCTCTCTCTCTCTCTCTCT 4779

QY 781 TAAACAACTTATTTCTATCTGACACAACTTTTATAGGCTATATCTATGTAATTTA 840
DB 4780 TAAACAACTTATTTCTATCTGACACAACTTTTATAGGCTATATCTATGTAATTTA 4839

QY 841 CTAGTATGTTGTTAGGATTTATGTTAAATTTGTTATATGTCGCGCCAAATCATTTCCAC 900
DB 4840 CTAGTATGTTGTTAGGATTTATGTTAAATTTGTTATATGTCGCGCCAAATCATTTCCAC 4899

QY 901 TAAATATGACATATATATCATAACTGCTTGGTGATAGCTCAGTGTTTAATAGTTATTCTC 960
DB 4900 TAAATATGACATATATATCATAACTGCTTGGTGATAGCTCAGTGTTTAATAGTTATTCTC 4959

QY 961 AGAAATCAAAATTTGATAGTTAAATACATTTATAGGCAAAATGCTAACTATT 1020
DB 4960 AGAAATCAAAATTTGATAGTTAAATACATTTATAGGCAAAATGCTAACTATT 5019

QY 1021 TCTACATTAATTTCAATTTTCCAGATAATGAATAATGAGTCAATGAGTACTCTCAGAACT 1080
DB 5020 TCTACATTAATTTCAATTTTCCAGATAATGAATAATGAGTCAATGAGTACTCTCAGAACT 5079

QY 1081 GGAAAGCACCAATTTATATATAGATGAGACTGTGAAATAGCAATATCCCAACTAACCTCG 1140
DB 5080 GGAAAGCACCAATTTATATATAGATGAGACTGTGAAATAGCAATATCCCAACTAACCTCG 5139

QY 1141 TGTGCTTCTGTTCAATCTCGGAAACCTGAGAAAGCAAAATACAAAAGTTAGAAATCTGATG 1200
DB 5140 TGTGCTTCTGTTCAATCTCGGAAACCTGAGAAAGCAAAATACAAAAGTTAGAAATCTGATG 5199

QY 1201 CTCAGCTCAAAATGGAATTTGTGCGACCCCATGCTGTGAGTGTGCAATATTTCTGTGTG 1260
DB 5200 CTCAGCTCAAAATGGAATTTGTGCGACCCCATGCTGTGAGTGTGCAATATTTCTGTGTG 5259

QY 1261 GTCTGCAAAAGTAACTGATTTCAATAAATATTTTGTAGAGGTTCCAGAAAGCACTCACAC 1320
DB 5260 GTCTGCAAAAGTAACTGATTTCAATAAATATTTTGTAGAGGTTCCAGAAAGCACTCACAC 5319

QY 1321 ACCAAAAATAGAGAAACCAACCAACAAAAATGCTAAAGTGGATTTTCCCAACAGATC 1380
DB 5320 ACCAAAAATAGAGAAACCAACCAACAAAAATGCTAAAGTGGATTTTCCCAACAGATC 5379

QY 1381 ATATGACATTAACAGTACATCAATAAATAATCTTGTGAGGCTGTGTTTGTGAGCTGGCT 1440
DB 5380 ATATGACATTAACAGTACATCAATAAATAATCTTGTGAGGCTGTGTTTGTGAGCTGGCT 5439

QY 1441 GGTGCAATTTGCTGTTTGTAGCAGAGATGGGCGACAGGTAGTCCCAGGGTGGCTGAT 1500
DB 5440 GGTGCAATTTGCTGTTTGTAGCAGAGATGGGCGACAGGTAGTCCCAGGGTGGCTGAT 5499

QY 1501 GTGTCATCTGCTTACTGCTTGAACAGATGCGAGAAACCAACAGATAGATGTAGAAGTTTC 1560
DB 5500 GTGTCATCTGCTTACTGCTTGAACAGATGCGAGAAACCAACAGATAGATGTAGAAGTTTC 5559

QY 1561 TCCATTTGTGTGTTCTGGAGCTCATGGAATATCCAGAGACACAAAGGTGAGAGAGC 1620
DB 5560 TCCATTTGTGTGTTCTGGAGCTCATGGAATATCCAGAGACACAAAGGTGAGAGAGC 5619

QY 1621 TTTGTTCTCTCTTGTAGCAGATTAACGCTCTCAAACTGGGTTGGACTTACTTAAAGTAA 1680
DB 5620 TTTGTTCTCTCTTGTAGCAGATTAACGCTCTCAAACTGGGTTGGACTTACTTAAAGTAA 5679

QY 1681 ATGAAATCTTAATATTTGTTATTTTCAAGGCTCTATAATAACACACTCTTCTAGTGA 1740
DB 5680 ATGAAATCTTAATATTTGTTATTTTCAAGGCTCTATAATAACACACTCTTCTAGTGA 5739

QY 1741 ACTTATGTAATGTTATTTTAAAGAAATTTGTGACTAAATAACAAAGTAAATTTATGCTAAAC 1800
DB 5740 ACTTATGTAATGTTATTTTAAAGAAATTTGTGACTAAATAACAAAGTAAATTTATGCTAAAC 5799

QY 1801 CCTCGAATATGTTGTTTACATTTTCAGAGATGAGGAAATTTATCAGGAAAGGAGGT 1860
DB 5800 CCTCGAATATGTTGTTTACATTTTCAGAGATGAGGAAATTTATCAGGAAAGGAGGT 5859

QY 1861 GAAACATCTGAAATGTTCTCTCAATCTCAACCTGACAGTCTGTCAAAACCGGTATAGAGTATAC 1920
DB 5860 GAAACATCTGAAATGTTCTCTCAATCTCAACCTGACAGTCTGTCTCAAAACCGGTATAGAGTATAC 5919

QY 1921 TGTGACATGAATACAGAAATAGAGGTAAAGCTTTTCGACAGTTGTTGACCTGTTGATCTGT 1980
DB 5920 TGTGACATGAATACAGAAATAGAGGTAAAGCTTTTCGACAGTTGTTGACCTGTTGATCTGT 5979

QY 1981 AATTATTTGGATACCGTAAATGCGGAAACCAAGCCAGGTGTGCTGGCTCATACCTGT 2040

Db 5980 AATTATTTGGATACCGTAAATGCCAGGAAACAGGCCAGGTGTGGTGGCTCATACCTGT 6039
QY 2041 AATTCACGACCTTGGGAGGCCAAAGTGGCTGTAGCTTGAGCCTAGAGTTTCAAAC 2100
Db 6040 AATTCACGACCTTGGGAGGCCAAAGTGGCTGTAGCTTGAGCCTAGAGTTTCAAAC 6099
QY 2101 AGCTGGGCAACATAATAGAGCCCTAACTCTACAAAATAAATAAATAAATAAATAA 2160
Db 6100 AGCTGGGCAACATAATAGAGCCCTAACTCTACAAAATAAATAAATAAATAAATAA 6159
QY 2161 AAAAAATCAGCTGTGGTGTAGTATGTGCTGTAGTCCAGCTATCCAGAGGCTGAGA 2220
Db 6160 AAAAAATCAGCTGTGGTGTAGTATGTGCTGTAGTCCAGCTATCCAGAGGCTGAGA 6219
QY 2221 TGGGAGATCACCTGAGGCCCAACCTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGG 2280
Db 6220 TGGGAGATCACCTGAGGCCCAACCTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGG 6279
QY 2281 GCAACAGAGGATGAGTGTAGATCCTGTCTCAAAAAATAAATAAATAAATAAATAA 2340
Db 6280 GCAACAGAGGATGAGTGTAGATCCTGTCTCAAAAAATAAATAAATAAATAAATAA 6339
QY 2341 AAGACTTAGCTCTAACATCTAACATGCTGACAAAGAGTAAATTTGATGTGGAATTCAC 2400
Db 6340 AAGACTTAGCTCTAACATCTAACATGCTGACAAAGAGTAAATTTGATGTGGAATTCAC 6399
QY 2401 CTGATATTTAAAGTTATAAATAATCTAATTTCAAAATTTGGGGTAAAGTAAAGCACTT 2460
Db 6400 CTGATATTTAAAGTTATAAATAATCTAATTTCAAAATTTGGGGTAAAGTAAAGCACTT 6459
QY 2461 GCAGTTTCCAAAGATTTCAAGTTTACCTCTCATATTTATTTTCTTATTTGTGTCTATT 2520
Db 6460 GCAGTTTCCAAAGATTTCAAGTTTACCTCTCATATTTATTTTCTTATTTGTGTCTATT 6519
QY 2521 TAGAGCACCAATATATATAATGGAATGGACAGGGATTCAGATATATTTTCAAAGT 2580
Db 6520 TAGAGCACCAATATATATAATGGAATGGACAGGGATTCAGATATATTTTCAAAGT 6579
QY 2581 GACATATTTGTGTGTGTAAATATATGCTCTTTTTTGTCTCTGCAACCAAGAGTGGAC 2640
Db 6580 GACATATTTGTGTGTGTAAATATATGCTCTTTTTTGTCTCTGCAACCAAGAGTGGAC 6639
QY 2641 AGTGATTCAGAACCTGCAAGCGGTGTTGACTTTTGGCAGGAAATGGATCCATATAA 2700
Db 6640 AGTGATTCAGAACCTGCAAGCGGTGTTGACTTTTGGCAGGAAATGGATCCATATAA 6699
QY 2701 ACAGGGATTTGGAAATGTTGCAACCAACACAGATGGGAAGATTTACTGTGGCTTACCAG 2760
Db 6700 ACAGGGATTTGGAAATGTTGCAACCAACACAGATGGGAAGATTTACTGTGGCTTACCAG 6759
QY 2761 TAAAGCAAGGATGCAAAATAAATAATCATTTCTATTTGAAATGGGATTTTTTTTAAATAAA 2820
Db 6760 TAAAGCAAGGATGCAAAATAAATAATCATTTCTATTTGAAATGGGATTTTTTTTAAATAAA 6819
QY 2821 AAACATTCATTTGGAGCCCTGTTTTAGGAGTTTAAAGAGGTTTCCCTGCAAAAAATGT 2880
Db 6820 AAACATTCATTTGGAGCCCTGTTTTAGGAGTTTAAAGAGGTTTCCCTGCAAAAAATGT 6879
QY 2881 GGAAGCTAAAGATAGGGAAGGAGGAGGAGTGTGTTTGTGTTTCCCAAAATTTATTTTGGTG 2940
Db 6880 GGAAGCTAAAGATAGGGAAGGAGGAGGAGTGTGTTTGTGTTTCCCAAAATTTATTTTGGTG 6939
QY 2941 AGAGATTTTATTTTGTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 3000
Db 6940 AGAGATTTTATTTTGTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 6999
QY 3001 AGCTTACAGGATGGGACCCACAGAACTTTTGTGATGAAATGAGGAGCTGGAAAGAGACA 3060
Db 7000 AGCTTACAGGATGGGACCCACAGAACTTTTGTGATGAAATGAGGAGCTGGAAAGAGACA 7059
QY 3061 AAGTAAAGGCTCACTATGAGGATTCATCTGACAGATGACCAACCAATACCAAGTCT 3120

Db 7060 AAGTAAAGGCTCACTATGAGAGATTCACTGTACAGAAATGAAGCCAAATAACAGATCT 7119
QY 3121 CAGTGAAACAAATACAGAGGAAACAGCCGGTAAATGCCCTCATGGATGGAGCATCTCAGCTGA 3180
Db 7120 CAGTGAAACAAATACAGAGGAAACAGCCGGTAAATGCCCTCATGGATGGAGCATCTCAGCTGA 7179
QY 3181 TGGGAGAAAAACAGGACCATGACCATTCACAGCGCATGTTCTTACGACGATATGACAGAG 3240
Db 7180 TGGGAGAAAAACAGGACCATGACCATTCACAGCGCATGTTCTTACGACGATATGACAGAG 7239
QY 3241 ACAATGACGGCTGTGTATGTGGCACTCTTTGCTCTCTTAAATAACACATTAATATC 3300
Db 7240 ACAATGACGGCTGTGTATGTGGCACTCTTTGCTCTCTTAAATAACACATTAATATC 7299
QY 3301 ATTACTCAGAATCATTAACAATATTTTAAATAGCTACCACTTCTCTGGGCATTTACTGTCA 3360
Db 7300 ATTACTCAGAATCATTAACAATATTTTAAATAGCTACCACTTCTCTGGGCATTTACTGTCA 7359
QY 3361 GCCACTGTCTTAAGTCTTTTATGCACTCTGAAAGCATTTTCAACTATAAGGTAGACATTT 3420
Db 7360 GCCACTGTCTTAAGTCTTTTATGCACTCTGAAAGCATTTTCAACTATAAGGTAGACATTT 7419
QY 3421 CTTATTTCTCATTTTACAGATGAGATTTAGAGAGATTTAGTGTGTTGTTCCAAATGTCACACA 3480
Db 7420 CTTATTTCTCATTTTACAGATGAGATTTAGAGAGATTTAGTGTGTTGTTCCAAATGTCACACA 7479
QY 3481 ACTACCCAGAGATAAACTAGAAATTTGAGCACAGTTACTTTCTGTAATTAATGAGCATTTAG 3540
Db 7480 ACTACCCAGAGATAAACTAGAAATTTGAGCACAGTTACTTTCTGTAATTAATGAGCATTTAG 7539
QY 3541 ATAAATACCTATATCTCTATATTTCTAAAGTGTGTGTGAAACCTTTTCAATTTTCAATTTCCAG 3600
Db 7540 ATAAATACCTATATCTCTATATTTCTAAAGTGTGTGTGAAACCTTTTCAATTTTCAATTTCCAG 7599
QY 3601 GCTTCTGTACTAAAGGTTGTAAAGCTATTTTCCAGTATAAGTAAACAAACACAGT 3660
Db 7600 GCTTCTGTACTAAAGGTTGTAAAGCTATTTTCCAGTATAAGTAAACAAACACAGT 7659
QY 3661 CCCTAGATGGATTTGCCACAAAGGCCAGTTATCTCTTTCTTGTATAGGGCACAGGAG 3720
Db 7660 CCCTAGATGGATTTGCCACAAAGGCCAGTTATCTCTTTCTTGTATAGGGCACAGGAG 7719
QY 3721 GTCTTTGTGTATTTAGTGTGACTCTATGTATAGCAACCAAGGAAAGACTACTGTGACACA 3780
Db 7720 GTCTTTGTGTATTTAGTGTGACTCTATGTATAGCAACCAAGGAAAGACTACTGTGACACA 7779
QY 3781 CGAGTGTAGCAGTCTTTTATGCGTAATCTGCAAAACGTAACCTTGACACCCGTTAGTTCTGT 3840
Db 7780 CGAGTGTAGCAGTCTTTTATGGTAACTCTGCAAAACGTAACCTTGACACCCGTTAGTTCTGT 7839
QY 3841 TTCTTAATAACGCCAAACACATTTTCTTTTCAAGTTTAACTCAGATCCCAGAAAACAGTGT 3900
Db 7840 TTCTTAATAACGCCAAACACATTTTCTTTTCAAGTTTAACTCAGATCCCAGAAAACAGTGT 7899
QY 3901 CTAAAGAAAGACGGTGTGGATGGTGTATATAGATGTATGCGACCAATCCAAACGGCA 3960
Db 7900 CTAAAGAAAGACGGTGTGGTGTATATAGATGTATGCGACCAATCCAAACGGCA 7959
QY 3961 GATACCTACTGGGTGGACAGTACCTCGGACATGCGCAAGCATGGCACAGATGATGTTG 4020
Db 7960 GATACCTACTGGGTGGACAGTACCTCGGACATGCGCAAGCATGGCACAGATGATGTTG 8019
QY 4021 TAGTATGGAATTTGGAAGGGGTCACTGTACTCAATGAGGAAGATGAGTATGAAGATCA 4080
Db 8020 TAGTATGGAATTTGGAAGGGGTCACTGTACTCAATGAGGAAGATGAGTATGAAGATCA 8079
QY 4081 GGCCCTTCTTCCACAGCAATAGTCCCAATAAGTATGTTTGTCTCTCTGTATGTGAC 4140
Db 8080 GGCCCTTCTTCCACAGCAATAGTCCCAATAAGTATGTTTGTCTCTCTGTATGTGAC 8139
QY 4141 AACATTTTGTGATCATTTATGGAATTTTCTTTTCAATATATATTTTCTCTTAAACT 4200
Db 8140 AACATTTTGTGATCATTTATGGAATTTTCTTTTCAATATATATTTTCTCTTAAACT 8199

[illegible]

Db	5500	GTGTGCATCTGCGTACTGGCTTTGAA	CAGATGGCAGAACCA	CAGATAGATCTGACAGTTTC	5555
QY	1561	TCCATTTTGTGTGTTCTGGAGCTCAT	TGGATATCCAGACACAAAAGTGG	AAGAGAGC	1620
Db	5560	TCCATTTTGTGTGTTCTGGAGCTCAT	TGGATATCCAGACACAAAAGTGG	AAGAGAGC	5619
QY	1621	TTTGTTCATCCTCTTAGCAGATAAA	CGTCCCAAACCTGGGTGGACTTACT	TAAGTAAA	1680
Db	5620	TTTGTTCATCCTCTTAGCAGATAAA	CGTCCCAAACCTGGGTGGACTTACT	TAAGTAAA	5679
QY	1681	ATGAAATCTAAATATTTGTTATATAT	TTTCAAAGGTCTATAATAACACAC	TCTCTTAGTA	1740
Db	5680	ATGAAATCTAAATATTTGTTATATAT	TTTCAAAGGTCTATAATAACACAC	TCTCTTAGTA	5739
QY	1741	ACTTATGTAAATGTTATTTTAAAGAA	TGGTGACCTAAATACAAAGTAAT	TATGTCTATAAC	1800
Db	5740	ACTTATGTAAATGTTATTTTAAAGAA	TGGTGACCTAAATACAAAGTAAT	TATGTCTATAAC	5799
QY	1801	CCCTGAAACATATGTGTCTTACATTT	TGCAGAAATGTAGGAAAATATCAG	GAAGAGGT	1860
Db	5800	CCCTGAAACATATGTGTCTTACATTT	TGCAGAAATGTAGGAAAATATCAG	GAAGAGGT	5859
QY	1861	GAAACATCTGAAATGTATCTCAATTC	CAACCTGACAGTTCTGTCAAAAC	CGTATAGAGTATAC	1920
Db	5860	GAAACATCTGAAATGTATCTCAATTC	CAACCTGACAGTTCTGTCAAAAC	CGTATAGAGTATAC	5919
QY	1921	TGTGACATGAATACAGAAAATGGAG	TAAGCTTTTCGACAGTTGTGTGAC	TCTTGTGATCTGT	1980
Db	5920	TGTGACATGAATACAGAAAATGGAG	TAAGCTTTTCGACAGTTGTGTGAC	TCTTGTGATCTGT	5979
QY	1981	AATTAATTTGGATACCGTFAAAATGC	CAGGAAACAAAGCCAGGTGTGGT	GGCTCATACCTGT	2040
Db	5980	AATTAATTTGGATACCGTFAAAATGC	CAGGAAACAAAGCCAGGTGTGGT	GGCTCATACCTGT	6039
QY	2041	AATTCAGACACCTTGGAGGCCAAAG	TGGGCTGATAGCTTGAGCCTAGG	AGTTTGAACCT	2100
Db	6040	AATTCAGACACCTTGGAGGCCAAAG	TGGGCTGATAGCTTGAGCCTAGG	AGTTTGAACCT	6099
QY	2101	AGCCTGGGCAACATAATGAGACCT	TAACCTCTACAAAATAAAATAAT	ACCAAAAAAA	2160
Db	6100	AGCCTGGGCAACATAATGAGACCT	TAACCTCTACAAAATAAAATAAT	ACCAAAAAAA	6159
QY	2161	AAAAAATCAGCTGTGTGTGTAGTAT	GTGCTGTAGTCCAGCTATCCAG	AGGCTGAGA	2220
Db	6160	AAAAAATCAGCTGTGTGTGTAGTAT	GTGCTGTAGTCCAGCTATCCAG	AGGCTGAGA	6219
QY	2221	TGGGAGATCACCTGAGCCCAACAA	CTGGAGTCTTGATCATGTCTGAA	CTGTAGCCTGG	2280
Db	6220	TGGGAGATCACCTGAGCCCAACAA	CTGGAGTCTTGATCATGTCTGAA	CTGTAGCCTGG	6279
QY	2281	GCAACAGAGGATAGTGAGATCTGT	CTCAAAAATAAAATAATTAATTA	AAAAAGCCAGAAAC	2340
Db	6280	GCAACAGAGGATAGTGAGATCTGT	CTCAAAAATAAAATAATTAATTA	AAAAAGCCAGAAAC	6339
QY	2341	AAGACTTAGCTCTAACATCTAA	CATAGCTGCAAAAGGAGTAATTT	GATGTGGAATTCAC	2400
Db	6340	AAGACTTAGCTCTAACATCTAA	CATAGCTGCAAAAGGAGTAATTT	GATGTGGAATTCAC	6399
QY	2401	CTGATATTTTAAAGTTATAAATAT	CTATAATTCACATTTGGGTAA	GATAAAGCACATT	2460
Db	6400	CTGATATTTTAAAGTTATAAATAT	CTATAATTCACATTTGGGTAA	GATAAAGCACATT	6459
QY	2461	GCAGTTTCCAAAGATTTTACAGTT	TACTCTCATATTTTCCCTTAT	TGTGCTATTT	2520
Db	6460	GCAGTTTCCAAAGATTTTACAGTT	TACTCTCATATTTTCCCTTAT	TGTGCTATTT	6519
QY	2521	TAGAGCCCAATATATACTAAAT	TGGAATGGACAGGGGATTCAG	ATATATTTTCAAAGT	2580
Db	6520	TAGAGCCCAATATATACTAAAT	TGGAATGGACAGGGGATTCAG	ATATATATTTTCAAAGT	6579
QY	2581	GACATTAATTTGCTGTTGGTTAA	TATATATGCTCTTTTGTCTG	CAACCAAGGATGAC	2640
Db	6580	GACATTAATTTGCTGTTGGTTAA	TATATATGCTCTTTTGTCTG	CAACCAAGGATGAC	6639

QY 2641 AGTGATTCAGAACCGTCAAGACGGTAGTGTGACCTTTGGCAGCAAAATGGGATCCCATATAA 2700
DB 6640 AGTGATTCAGAACCGTCAAGACGGTAGTGTGACCTTTGGCAGCAAAATGGGATCCCATATAA 6699
QY 2701 ACAGGGATTTGAAATGTTGCAACCAACACAGATGGGAAGAAATTAAGTGTGGCTACCCAGG 2760
DB 6700 ACAGGGATTTGAAATGTTGCAACCAACACAGATGGGAAGAAATTAAGTGTGGCTACCCAGG 6759
QY 2761 TAACGGAACAGGCATGCAAAATAAATAATCATTTCTATTGAAATGGGATTTTTTTTAAATATAA 2820
DB 6760 TAACGGAACAGGCATGCAAAATAAATAATCATTTCTATTGAAATGGGATTTTTTTTAAATATAA 6819
QY 2821 AAAACATTCAATGTTGGAACCGCTTTTAGGCAGTTTAAGAGGAGTTTCCTGACAAAATGCT 2880
DB 6820 AAAACATTCAATGTTGGAACCGCTTTTAGGCAGTTTAAGAGGAGTTTCCTGACAAAATGCT 6879
QY 2881 GGAAGCTAAAGATTAAGGGAAGAAAGGCAGTTTTTACTTTCCCAAAATTTTTTTTGGTG 2940
DB 6880 GGAAGCTAAAGATTAAGGGAAGAAAGGCAGTTTTTACTTTCCCAAAATTTTTTTTGGTG 6939
QY 2941 AGAGATTTTTATTGTTTTTCTTTTAGGTGAATATTGGCTTGGAAATGATAAATTAGCC 3000
DB 6940 AGAGATTTTTATTGTTTTTCTTTTAGGTGAATATTGGCTTGGAAATGATAAATTAGCC 6999
QY 3001 AGCTTACCAGATGGACCCACAGAACTTTTGTAGTAAGAAATGGAGACTGGAAAGGAGACA 3060
DB 7000 AGCTTACCAGATGGACCCACAGAACTTTTGTAGTAAGAAATGGAGACTGGAAAGGAGACA 7059
QY 3061 AAGTAAAGCTCACTATGAGAGATTCACATGTCAGAAATGAAGCAACAAATACAGATCT 3120
DB 7060 AAGTAAAGCTCACTATGAGAGATTCACATGTCAGAAATGAAGCAACAAATACAGATCT 7119
QY 3121 CAGTGAACAAATACAGAGGAACAGCGGTAAATGCCCTCATGAGATGGAGCAATCTCAGCTGA 3180
DB 7120 CAGTGAACAAATACAGAGGAACAGCGGTAAATGCCCTCATGAGATGGAGCAATCTCAGCTGA 7179
QY 3181 TGGGAGAAACAGGACCATGACATTCACAAACGGCATGTTCTTCAGCAGCTATGACAGAG 3240
DB 7180 TGGGAGAAACAGGACCATGACATTCACAAACGGCATGTTCTTCAGCAGCTATGACAGAG 7239
QY 3241 ACAATGACGCTGATGCTGTGACACTTTGCTCTGCTTTAAATAACACACTAATATC 3300
DB 7240 ACAATGACGCTGATGCTGTGACACTTTGCTCTGCTTTAAATAACACACTAATATC 7299
QY 3301 ATTACTCAGAAATCATTAACAATAATTTTAAATAGCTACCACTTCTGCGCACTTACTGTCA 3360
DB 7300 ATTACTCAGAAATCATTAACAATAATTTTAAATAGCTACCACTTCTGCGCACTTACTGTCA 7359
QY 3361 GCCACTGTCTAAGCTCTTTATGCAATCACTCGAAAGCAATTTCAACTATAGGTAGACAT 3420
DB 7360 GCCACTGTCTAAGCTCTTTATGCAATCACTCGAAAGCAATTTCAACTATAGGTAGACAT 7419
QY 3421 CTTATTCTCAATTTACAGATGAGATTTAGAGAGATTACGTGATTTGCCAATGTCAACA 3480
DB 7420 CTTATTCTCAATTTACAGATGAGATTTAGAGAGATTACGTGATTTGCCAATGTCAACA 7479
QY 3481 ACTACCAGAGATAAACTAGAATTTGAGCAACAGTTACTTTCTGAAATAATGAGCATTTAG 3540
DB 7480 ACTACCAGAGATAAACTAGAATTTGAGCAACAGTTACTTTCTGAAATAATGAGCATTTAG 7539
QY 3541 ATAATACTTATCTCTATATTCTAAAGTGTGTGAAACTTTTCATTTTCCAG 3600
DB 7540 ATAATACTTATCTCTATATTCTAAAGTGTGTGAAACTTTTCATTTTCCAG 7599
QY 3601 GGTTCCTGATACCTAAGGGTTGTAAGCTAATTTCCAGTATAAAGTAAACAAACACAGT 3660
DB 7600 GGTTCCTGATACCTAAGGGTTGTAAGCTAATTTCCAGTATAAAGTAAACAAACACAGT 7659
QY 3661 CCCTAGATGGATTGCCAACAAGGCCAGTTATCTCTCTTTCTGCTATAGGGCAGGAG 3720
DB 7660 CCCTAGATGGATTGCCAACAAGGCCAGTTATCTCTCTTTCTGCTATAGGGCAGGAG 7719

QY 3721 GTCTTTGGTGTATTAGTGTGACTCTATGTATAGCACCCAAAGAAAGACTACTGTGCACA 3780
DB 7720 GTCTTTGGTGTATTAGTGTGACTCTATGTATAGCACCCAAAGAAAGACTACTGTGCACA 7779
QY 3781 CGAGTGTAGCAGTCTTTTATGGGTAATCTGMAAACGTAACTTTGACCAACCGTAGTCTCTG 3840
DB 7780 CGAGTGTAGCAGTCTTTTATGGGTAACTCTGCAAAAACGTAACTTTGACCAACCGTAGTCTCTG 7839
QY 3841 TTCTAAATAACGCCAAACACATTTCTTTTCAGGTAAACATCAGATCCCAAGAAACAGTGT 3900
DB 7840 TTCTAAATAACGCCAAACACATTTCTTTTCAGGTAAACATCAGATCCCAAGAAACAGTGT 7899
QY 3901 CTAAGAAGACCGTGTGGATGGTATGATTAATAGATGTCATCCAGCAATCCAAACGGCA 3960
DB 7900 CTAAGAAGACCGTGTGGATGGTATGATTAATAGATGTCATCCAGCAATCCAAACGGCA 7959
QY 3961 GATACCTACTGGGGTGGACAGTACACTGGGACATGGCAAGCATGGCAGAGATGATGGTG 4020
DB 7960 GATACCTACTGGGGTGGACAGTACACTGGGACATGGCAAGCATGGCAGAGATGATGGTG 8019
QY 4021 TAGTATGGATGAATTTGGAAGGGGTCACTGTACTCAATGAGGAAGATGAGTATGAAGATCA 4080
DB 8020 TAGTATGGATGAATTTGGAAGGGGTCACTGTACTCAATGAGGAAGATGAGTATGAAGATCA 8079
QY 4081 GGCCCTCTTCCACAGCAATAGTCCCAATACGTAGATTTTGTCTCTCTGATGTGAC 4140
DB 8080 GGCCCTCTTCCACAGCAATAGTCCCAATACGTAGATTTTGTCTCTCTGATGTGAC 8139
QY 4141 AACATTTTGTACATTTATGTTATTTGGAATTTCTTTTCATACATTTATTTTCTTAAACT 4200
DB 8140 AACATTTTGTACATTTATGTTATTTGGAATTTCTTTTCATACATTTATTTTCTTAAACT 8199
QY 4201 CTCACGACAGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAAATTAATTTAAATAG 4260
DB 8200 CTCACGACAGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAAATTAATTTAAATAG 8259
QY 4261 CACATGATTTCTTTTGTCTTTCTTCATTTCTCTGCTCAACCAAGAAAGTAAACAAAGTAT 4320
DB 8260 CACATGATTTCTTTTGTCTTTCTTCATTTCTCTGCTCAACCAAGAAAGTAAACAAAGTAT 8319
QY 4321 AGTTTTGACAGAGTGTGGTGTTCATTTTTCAGTTCTAGTTGATTCGAGAAATTTTCAAT 4380
DB 8320 AGTTTTGACAGAGTGTGGTGTTCATTTTTCAGTTCTAGTTGATTCGAGAAATTTTCAAT 8379
QY 4381 AAGGAGAGGGGTCTTTTATCTGTCGAGGAAACCATGACGGAAGGAAACAACTGAT 4440
DB 8380 AAGGAGAGGGGTCTTTTATCTGTCGAGGAAACCATGACGGAAGGAAACAACTGAT 8439
QY 4441 GTTTAAAGTCCACTTTTAAACTATATTTTATTTATGTTAGGATCTGTCAAAGAAACCTTC 4500
DB 8440 GTTTAAAGTCCACTTTTAAACTATATTTTATTTATGTTAGGATCTGTCAAAGAAACCTTC 8499
QY 4501 CAAAAAGATTTTATTAATTAACCCAGACTCTGTTGCAATAGTTAATGTTTCTTGTGTTTG 4560
DB 8500 CAAAAAGATTTTATTAATTAACCCAGACTCTGTTGCAATAGTTAATGTTTCTTGTGTTTG 8559
QY 4561 TAATCCACACATTTCAATGAGTGTAGGCTTTGCACCTTTGAAGGAAGGAGAGCTTCACAAC 4620
DB 8560 TAATCCACACATTTCAATGAGTGTAGGCTTTGCACCTTTGAAGGAAGGAGAGCTTCACAAC 8619
QY 4621 CTCAATAGCTAATAAACCCGCTCTTGAATATTTGAAGATTTTAAATCTGACTCTAGGAGC 4680
DB 8620 CTCAATAGCTAATAAACCCGCTCTTGAATATTTGAAGATTTTAAATCTGACTCTAGGAGC 8679
QY 4681 GGCAACGCTGTCTACGACTATAATCCCAACAATTTTGGAGGCTGAGCGGGCGGTCAAA 4740
DB 8680 GGCAACGCTGTCTACGACTATAATCCCAACAATTTTGGAGGCTGAGCGGGCGGTCAAA 8739
QY 4741 GGTGAGAGTTCAAGACAGGCTGACCAATATGTTGAACCCCACTCTACTATAAATATAC 4800
DB 8740 GGTGAGAGTTCAAGACAGGCTGACCAATATGTTGAACCCCACTCTACTATAAATATAC 8799
QY 4801 AAAAAATTAGCAGCGGTGGTGGCAGGTGCCTGTAGGTCCAGCTAGCCTGTGAGGTGGAG 4860

Db 8800 AAAAAATAGCCAGCGTGGGAGGTCCTGTAGGTCCTCCAGCTAGCTGTGAGTGGAG 8859
QY 4861 ATTGCATTGAGCCCAAGATC 4879
Db 8860 ATTGCATTGAGCCCAAGATC 8878
RESULT 4
ADP65553
ID ADP65553 standard; DNA; 8878 BP.
AC ADP65553;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human fibrinogen beta chain (FGB), complete cds DNA.
XX
KW autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; angiot; antinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; Osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; ds; human.
XX
OS Homo sapiens.
XX
XX WO2003072827-A1.
PN
XX
PD 04-SEP-2003.
XX
PF 31-OCT-2002; 2002WO-US035433.
XX
XX 31-OCT-2001; 2001US-0336220P.
PR
XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX
XX Hirsch R, Thorton SL;
PI
XX WPI; 2003-712740/67.
DR
XX GENBANK; M64983.
XX
PT Diagnosing and analyzing autoimmune disease using gene expression
PT profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
XX
PS Disclosure; Page; 56pp; English.
XX
CC The invention relates to a novel method for diagnosing and analysing
CC autoimmune disease or arthritides. The method comprises obtaining a
CC patient sample containing mRNA, analysing gene expression using the mRNA
CC that results in a gene expression signature of the mRNA, and using that
CC gene expression signature to diagnose or analyse the autoimmune disease
CC or arthritides in the patient, where gene expression of at least 60% of
CC the genes correlates with that of the gene signature. The invention
CC further comprises: a treatment of rheumatoid arthritis; identification of
CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC analyses of autoimmune disease or rheumatoid arthritis; screening the
CC efficacy of a candidate drug in vitro for the treatment of collagen-
CC induced arthritis; and reducing the symptoms associated with collagen-
CC induced arthritis. The compositions of the invention have the following
CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC angiot, antinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This polynucleotide
CC represents a DNA sequence relating to the genes used in the analysis and
CC treatment of autoimmune diseases or arthritides. Note: This sequence is

CC not shown in the specification. It has been supplied in an electronic
CC format from WIFO.
XX
SQ Sequence 8878 BP; 2912 A; 1478 C; 1663 G; 2825 T; 0 U; 0 Other;
Query Match 100.0%; Score 4879; DB 11; Length 8878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCCGAGACCTCTCTT 60
Db 4000 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCCGAGACCTCTCTT 4059
QY 61 CTTCCCTTCAGTACATGTAATTTGCTGAAAGACCTGTGCAAAAGAGGCGAGCAAGTAA 120
Db 4060 CTTCCCTTCAGTACATGTAATTTGCTGAAAGACCTGTGCAAAAGAGGCGAGCAAGTAA 4119
QY 121 AAGGTAGATATCTGTGTGCTTCCATTCCATTTCAGCTTATAAAATTTGGAACCGTTAGAC 180
Db 4120 AAGGTAGATATCTGTGTGCTTCCATTTCAGCTTATAAAATTTGGAACCGTTAGAC 4179
QY 181 TGCCACGAGAATGCATGTTGTGAGAGATTAAACATTTCTGGGTAGTGAATGACATTC 240
Db 4180 TGCCACGAGAATGCATGTTGTGAGAGATTAAACATTTCTGGGTAGTGAATGACATTC 4239
QY 241 TACGCTTTTGGGCACCTTCCCTGCAACTTGCAGATAGCAGCTATTTCAGCTCTTATTC 300
Db 4240 TACGCTTTTGGGCACCTTCCCTGCAACTTGCAGATAGCAGCTATTTCAGCTCTTATTC 4299
QY 301 CAGTCTGACATCAGCAAGTGTGATTTCTATGAAAAATTTCTACTATGATCTCTTATTTA 360
Db 4300 CAGTCTGACATCAGCAAGTGTGATTTCTATGAAAAATTTCTACTATGATCTCTTATTTA 4359
QY 361 AGTATACAGAACTTTGTGACTCAGAGATAATATTACAGAGTGGAAAAAACCCTTAG 420
Db 4360 AGTATACAGAACTTTGTGACTCAGAGATAATATTACAGAGTGGAAAAAACCCTTAG 4419
QY 421 CATTTATAGTTTAAACATTTGAGGTTTGAATGAGAGATTTATCATATATATATTTCA 480
Db 4420 CATTTATAGTTTAAACATTTGAGGTTTGAATGAGAGATTTATCATATATATATTTCA 4479
QY 481 GTGTGTGGATAATGACACCTAACTGTGAATTTTGTAGGTGAGAATTTGTAGTGTGTTG 540
Db 4480 GTGTGTGGATAATGACACCTAACTGTGAATTTTGTAGGTGAGAATTTGTAGTGTGTTG 4539
QY 541 ACTTGGTGTGAGGAAACAGCTAGTGGCTGAGCCTGGGCACAGGCACTCTCAGTGTAGCA 600
Db 4540 ACTTGGTGTGAGGAAACAGCTAGTGGCTGAGCCTGGGCACAGGCACTCTCAGTGTAGCA 4599
QY 601 TACCCACAGTTGGAAAAATTTTCAAAGAAATCAAAGGAATCATGACATCTTATAAAATTTCA 660
Db 4600 TACCCACAGTTGGAAAAATTTTCAAAGAAATCAAAGGAATCATGACATCTTATAAAATTTCA 4659
QY 661 AGGTTCTGCTATATCTATGTGAATTTGATAAATAAATCAAGCATATCCACTCTGTGAAT 720
Db 4660 AGGTTCTGCTATATCTATGTGAATTTGATAAATAAATCAAGCATATCCACTCTGTGAAT 4719
QY 721 TGAATCTCTCAGATGGAGACCCCAATCTGTTCTCTCTTTTCCCTCACCAGAAAA 780
Db 4720 TGAATCTCTCAGATGGAGACCCCAATCTGTTCTCTCTTTTCCCTCACCAGAAAA 4779
QY 781 TAAACAACCTATTTTCAATTTTACTTGGACACAAATCTTTAGCGTATACCTATGGTAAATTA 840
Db 4780 TAAACAACCTATTTTCAATTTTACTTGGACACAAATCTTTAGCGTATACCTATGGTAAATTA 4839
QY 841 CTAGTATGGTGTAGGATTTATGTTAATTTGTATATGTCTATGCGCCCAAAATCATTTCCAC 900
Db 4840 CTAGTATGGTGTAGGATTTATGTTAATTTGTATATGTCTATGCGCCCAAAATCATTTCCAC 4899
QY 901 TAAATATGACTAT 960
Db 4900 TAAATATGACTAT 4959

QY 961 AGAATCAAAATTTGATAGTTAAATACATTAAGTTTATGAGCCAAAATGCTAACTATT 1020
DB |||||
QY 4960 AGAAATCAAAATTTGATAGTTAAATACATTAAGTTTATGAGCCAAAATGCTAACTATT 5019
DB |||||
QY 1021 TCTACATAATTTCAATTTTCCAGATAATGAAATAGTCAATGAGTACTCTCGAAGT 1080
DB |||||
QY 5020 TCTACATAATTTCAATTTTCCAGATAATGAAATAGTCAATGAGTACTCTCGAAGT 5079
DB |||||
QY 1081 GGAAAGCAACCAATTTATATATAGATGAGACTGTGAATAGCAATATCCCAACTAACTCTTCG 1140
DB |||||
QY 5080 GGAAAGCAACCAATTTATATATAGATGAGACTGTGAATAGCAATATCCCAACTAACTCTTCG 5139
DB |||||
QY 1141 TGTGCTTCTGCTCAATCTCGAAGACTGTGAGAGCAAAATACAAAAGTTAGAAATCTCATGT 1200
DB |||||
QY 5140 TGTGCTTCTGCTCAATCTCGAAGACTGTGAGAGCAAAATACAAAAGTTAGAAATCTCATGT 5199
DB |||||
QY 1201 CTCAGCTCAAAATGGAATATTTGTCGACCCCATGCACTGTCAAGTGAATATTTCTGTGTT 1260
DB |||||
QY 5200 CTCAGCTCAAAATGGAATATTTGTCGACCCCATGCACTGTCAAGTGAATATTTCTGTGTT 5259
DB |||||
QY 1261 GTCTGGCAAGGTAATCTGAATCTAAACATATTTTTAGAGAGTTCCAGAGAACTCACAC 1320
DB |||||
QY 5260 GTCTGGCAAGGTAATCTGAATCTAAACATATTTTTAGAGAGTTCCAGAGAACTCACAC 5319
DB |||||
QY 1321 ACCAAAATAAGAGAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1380
DB |||||
QY 5320 ACCAAAATAAGAGAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 5379
DB |||||
QY 1381 ATAAATGACATTAACAGTACATCAATAAAATATCTTAGCCAGTTGTGTTTGGACTGGCCT 1440
DB |||||
QY 5380 ATAAATGACATTAACAGTACATCAATAAAATATCTTAGCCAGTTGTGTTTGGACTGGCCT 5439
DB |||||
QY 1441 GGTGCATTTGCTGGTTTGTAGAGAGGATGGGGCACAGGTAGTCCAGGGGTGGCTGAT 1500
DB |||||
QY 5440 GGTGCATTTGCTGGTTTGTAGAGAGGATGGGGCACAGGTAGTCCAGGGGTGGCTGAT 5499
DB |||||
QY 1501 GTGTGCATCTGCTAGTCTGCTGCAAGATGGCAGAACACAGATAGATGTAGAAGTTTC 1560
DB |||||
QY 5500 GTGTGCATCTGCTAGTCTGCTGCAAGATGGCAGAACACAGATAGATGTAGAAGTTTC 5559
DB |||||
QY 1561 TCCATTTTGTGTTCTGGAGTCTATGGATATTTCCAGGACACAAAAGGTGGAGAAGAC 1620
DB |||||
QY 5560 TCCATTTTGTGTTCTGGAGTCTATGGATATTTCCAGGACACAAAAGGTGGAGAAGAC 5619
DB |||||
QY 1621 TTTGTTCACTCTTAGCAGATAAACGCTCTCAAACTGGGTTGCACTTACTAAAGTAAA 1680
DB |||||
QY 5620 TTTGTTCACTCTTAGCAGATAAACGCTCTCAAACTGGGTTGCACTTACTAAAGTAAA 5679
DB |||||
QY 1681 ATGAAATCTAATATTTGTTATATTTTCAAGGCTATATAACACACTCTCTTAGTA 1740
DB |||||
QY 5680 ATGAAATCTAATATTTGTTATATTTTCAAGGCTATATAACACACTCTCTTAGTA 5739
DB |||||
QY 1741 ACTTATGTAATTTATTTTAAAGAAATTTGGTGAATAAATACAAAAGTAAATATGTCATAAAC 1800
DB |||||
QY 5740 ACTTATGTAATTTATTTTAAAGAAATTTGGTGAATAAATACAAAAGTAAATATGTCATAAAC 5799
DB |||||
QY 1801 CCTGGAACATAATGTTGCTTACATTTGCGAGATGTGAGGAATTTATCAGAAAGGAGGT 1860
DB |||||
QY 5800 CCTGGAACATAATGTTGCTTACATTTGCGAGATGTGAGGAATTTATCAGAAAGGAGGT 5859
DB |||||
QY 1861 GAAACATCTGAAATGTAATCTCAATTCACCTGACAGTTCTGTGCAAAACGTTATAGAGTATAC 1920
DB |||||
QY 5860 GAAACATCTGAAATGTAATCTCAATTCACCTGACAGTTCTGTGCAAAACGTTATAGAGTATAC 5919
DB |||||
QY 1921 TGTGACATGAATACAGAAAATGGAGGTAAAGTTTTCGACAGTTGTGACCTGTTGATCTGT 1980
DB |||||
QY 5920 TGTGACATGAATACAGAAAATGGAGGTAAAGTTTTCGACAGTTGTGACCTGTTGATCTGT 5979
DB |||||
QY 1981 AATTTATTTGGATACCGTAAATGTCAGAAACAAAGGCCAGGTGGTGGCTCATACCTGT 2040
DB |||||
QY 5980 AATTTATTTGGATACCGTAAATGTCAGAAACAAAGGCCAGGTGGTGGCTCATACCTGT 6039
DB |||||
QY 2041 AATTTCCAGCACCTTTGGGAGGCCAAAGTGGGCTGATAGCTTAGGCTTAGGAGTTTGAAC 2100
DB |||||

DB 6040 AATTTCCAGCACCTTTGGAGGCCAAAAGTGGGCTGTAGCTTGAGCCTTAGGAGTTTGAAC 6099
QY |||||
DB 2101 AGCCTGGGCAACATTAATGAGACCTTAACTTCAAAAAAATAAATAAATAAATAAATAA 2160
QY |||||
DB 6100 AGCCTGGGCAACATTAATGAGACCTTAACTTCAAAAAAATAAATAAATAAATAAATAA 6159
QY |||||
QY 2161 AAAAAATCAGCTGTGTTGGTAGTATGCTGTAGTCCAGCTATCCAGGAGGCTGAGA 2220
DB |||||
QY 6160 AAAAAATCAGCTGTGTTGGTAGTATGCTGTAGTCCAGCTATCCAGGAGGCTGAGA 6219
DB |||||
QY 2221 TGGGAGATCACCTGAGCCACCAACCTGGAGTCTTGATCATGCTAGTAACTGTAGCCTGG 2280
DB |||||
QY 6220 TGGGAGATCACCTGAGCCACCAACCTGGAGTCTTGATCATGCTAGTAACTGTAGCCTGG 6279
DB |||||
QY 2281 GCAACAGAGGATAGTGAGATCTGTCTCAAAAAAATAAATAAATAAATAAATAAATAA 2340
DB |||||
QY 6280 GCAACAGAGGATAGTGAGATCTGTCTCAAAAAAATAAATAAATAAATAAATAAATAA 6339
DB |||||
QY 2341 AAGACTTAGCTCTTAACATCTAATAGCTGACAAAGGAGTAAATTTGATGTGGAATTCAC 2400
DB |||||
QY 6340 AAGACTTAGCTCTTAACATCTAATAGCTGACAAAGGAGTAAATTTGATGTGGAATTCAC 6399
DB |||||
QY 2401 CTGATATTTTAAAGTTTATAAATAATCTATAATTTCAAAATTTGGGGTAAAGTAAAGCACTT 2460
DB |||||
QY 6400 CTGATATTTTAAAGTTTATAAATAATCTATAATTTCAAAATTTGGGGTAAAGTAAAGCACTT 6459
DB |||||
QY 2461 GCGATTTCCAAAGATTTTACAAGTTTACCTCTCATATTTATTTCTTATTTGCTCTATTT 2520
DB |||||
QY 6460 GCGATTTCCAAAGATTTTACAAGTTTACCTCTCATATTTATTTCTTATTTGCTCTATTT 6519
DB |||||
QY 2521 TAGAGCACCAAAATATATATCTAAATGGAATGGACAGGGGATTCAGATATATTTTCAAAAGT 2580
DB |||||
QY 6520 TAGAGCACCAAAATATATATCTAAATGGAATGGACAGGGGATTCAGATATATTTTCAAAAGT 6579
DB |||||
QY 2581 GACATATTTTGTGTTGGTTAAATATATATGCTCTTTTTGTTTCTGTCACAAAGGATGGAC 2640
DB |||||
QY 6580 GACATATTTTGTGTTGGTTAAATATATATGCTCTTTTTGTTTCTGTCACAAAGGATGGAC 6639
DB |||||
QY 2641 AGTGATTCGAACCGTCAACGAGTGTGATGTTGAGCTTTGGCAGGAAATGGATCCATATA 2700
DB |||||
QY 6640 AGTGATTCGAACCGTCAACGAGTGTGATGTTGAGCTTTGGCAGGAAATGGATCCATATA 6699
DB |||||
QY 2701 ACAGGATTTGGAATTTGTCACCAACACAGATGGGAATTTACTGTGGCCTACCAGG 2760
DB |||||
QY 6700 ACAGGATTTGGAATTTGTCACCAACACAGATGGGAATTTACTGTGGCCTACCAGG 6759
DB |||||
QY 2761 TAACGAACAGGCAATGCAAAATAAATAATCATTTCTATTTGAAATGGGATTTTTTTTAAATPAAA 2820
DB |||||
QY 6760 TAACGAACAGGCAATGCAAAATAAATAATCATTTCTATTTGAAATGGGATTTTTTTTAAATPAAA 6819
DB |||||
QY 2821 AAACATTTCTGTTGGAAGCTGTTTTAGGAGTAAAGGAGTTCCTGACAAAATGT 2880
DB |||||
QY 6820 AAACATTTCTGTTGGAAGCTGTTTTAGGAGTAAAGGAGTTCCTGACAAAATGT 6879
DB |||||
QY 2881 GGAAGCTAAAGATAAGGGAAGGAGGAGGTTTTAGTTTCCCAAAATTTTTTATTTTGGTG 2940
DB |||||
QY 6880 GGAAGCTAAAGATAAGGGAAGGAGGAGGTTTTAGTTTCCCAAAATTTTTTATTTTGGTG 6939
DB |||||
QY 2941 AGAGATTTTATTTTGTGTTTTCTTTTAGGTGAAATATTTGGCTTGGAAATGATAAATATGACC 3000
DB |||||
QY 6940 AGAGATTTTATTTTGTGTTTTCTTTTAGGTGAAATATTTGGCTTGGAAATGATAAATATGACC 6999
DB |||||
QY 3001 AGCTTACAGGATGGGACCCACAGAACTTTTGATAGAAATGGAGGACTGGAAGGAGACACA 3060
DB |||||
QY 7000 AGCTTACAGGATGGGACCCACAGAACTTTTGATAGAAATGGAGGACTGGAAGGAGACACA 7059
DB |||||
QY 3061 AAGTAAAGGCTCACTATGAGAGGATTCAGTCTACAGATGAAGCCAAACAAATACAGATCT 3120
DB |||||
QY 7060 AAGTAAAGGCTCACTATGAGAGGATTCAGTCTACAGATGAAGCCAAACAAATACAGATCT 7119
DB |||||
QY 3121 CAGTGAACAAATACAGAGGAACAGCCGTTAATGCCCTCATGGATGGAGCATCTCAGCTGA 3180
DB |||||

Db 7120 CAGTGAACAATACAGAGAAACAGCGGTAATGCCCTCATGGATGGAGCAATCTCAGCTGA 7179
QY 3181 TGGGAGAAACAGGACCATGACCAATTCACAAACGGCATGTTCTTCAGCACGATATGACAGAG 3240
Db 7180 TGGGAGAAACAGGACCATGACCAATTCACAAACGGCATGTTCTTCAGCACGATATGACAGAG 7239
QY 3241 ACAATGACGGCTGGTATGTTGGCACTCTTTGCTCTGCTTTAAAAATCACATAATC 3300
Db 7240 ACAATGACGGCTGGTATGTTGGCACTCTTTGCTCTGCTTTAAAAATCACATAATC 7299
QY 3301 ATTACTCAGAAATCAATTAACAATATTTTAAATAGCTACCACTCTCGGGCACTTACTGTCA 3360
Db 7300 ATTACTCAGAAATCAATTAACAATATTTTAAATAGCTACCACTCTCGGGCACTTACTGTCA 7359
QY 3361 GCCACTGTCTTAAGCTCTTTATGCACTCACTCGAAAGCAATTCACATTAAGGTAGACATT 3420
Db 7360 GCCACTGTCTTAAGCTCTTTATGCACTCACTCGAAAGCAATTCACATTAAGGTAGACATT 7419
QY 3421 CTATTTCTCATTTTACAGATGAGATTTAGAGATTAAGTGTCTTGAATTAATGAGCAATTTAG 3480
Db 7420 CTATTTCTCATTTTACAGATGAGATTTAGAGATTAAGTGTCTTGAATTAATGAGCAATTTAG 7479
QY 3481 ACTACCCAGAGATATAAATAGAAATTTAGCACAGTTTACTTCTGTAATTAATGAGCAATTTAG 3540
Db 7480 ACTACCCAGAGATATAAATAGAAATTTAGCACAGTTTACTTCTGTAATTAATGAGCAATTTAG 7539
QY 3541 ATAAATACCTATCTATCTATATCTTAAAGTGTGTGAAACCTTTCAATTTTCAATTTCCAG 3600
Db 7540 ATAAATACCTATCTATCTATATCTTAAAGTGTGTGAAACCTTTCAATTTTCAATTTCCAG 7599
QY 3601 GGTTCCTCATCTAAGGCTTCTTAAAGCTATTTTCCAGTATTAATGAGTAAAGCAATTTAG 3660
Db 7600 GGTTCCTCATCTAAGGCTTCTTAAAGCTATTTTCCAGTATTAATGAGTAAAGCAATTTAG 7659
QY 3661 CCTAGATGGAATGACAAAGCCAGTTATCTCTCTTCTTCTGCTATAGGCGACAGGAG 3720
Db 7660 CCTAGATGGAATGACAAAGCCAGTTATCTCTCTTCTTCTGCTATAGGCGACAGGAG 7719
QY 3721 GTCTTTGGTATTAAGTGTGACTCTATGATATAGCAACCCAAAGAAAGACTACTGTGCACA 3780
Db 7720 GTCTTTGGTATTAAGTGTGACTCTATGATATAGCAACCCAAAGAAAGACTACTGTGCACA 7779
QY 3781 CGAGTGTAGCACTCTTTTATGGTAACTCTGCAAAAGCTAACTTGACCCACCCGTTCTGT 3840
Db 7780 CGAGTGTAGCACTCTTTTATGGTAACTCTGCAAAAGCTAACTTGACCCACCCGTTCTGT 7839
QY 3841 TTCTTAATAACGCAAAACACATTTTCTTTCAGGTAAACATCAGATCCAGAAACAGAGTTT 3900
Db 7840 TTCTTAATAACGCAAAACACATTTTCTTTCAGGTAAACATCAGATCCAGAAACAGAGTTT 7899
QY 3901 CTAAAGAAAGACGGTGGTGGATGGTATTAATAGATGTATGATGATCCAGCCAAATCCAAACGGCA 3960
Db 7900 CTAAAGAAAGACGGTGGTGGTGGTATTAATAGATGTATGATGATCCAGCCAAATCCAAACGGCA 7959
QY 3961 GATACCTACTGGGCTGGACAGTACACCTGGACATGCGCAAGCATGGCACAGATGATGTTG 4020
Db 7960 GATACCTACTGGGCTGGACAGTACACCTGGACATGCGCAAGCATGGCACAGATGATGTTG 8019
QY 4021 TAGTATGATGAATTTGGAAGGGCTCATGGTACTCAATGAGGAAGATGAGTATGAAGATCA 4080
Db 8020 TAGTATGATGAATTTGGAAGGGCTCATGGTACTCAATGAGGAAGATGAGTATGAAGATCA 8079
QY 4081 GGCCCTCTTCCACAGCAATAGTCCCAATACGTFAGATTTTGTCTCTCTGTATGTGAC 4140
Db 8080 GGCCCTCTTCCACAGCAATAGTCCCAATACGTFAGATTTTGTCTCTCTGTATGTGAC 8139
QY 4141 AACATTTTGTATATGTTATGGAATTTTCTTTCATACATTAATTTCTCTAAACT 4200
Db 8140 AACATTTTGTATATGTTATGGAATTTTCTTTCATACATTAATTTCTCTAAACT 8199
QY 4201 CTCAGACGACGTGAGTGTGACTTTTGAAGAAAGTATAGGATAAATACATTTAAATAG 4260
Db 8200 CTCAGACGACGTGAGTGTGACTTTTGAAGAAAGTATAGGATAAATACATTTAAATAG 8259

QY 4261 CACATGATTTTCTTTTGTGTTTCTTCAATTTCTCTTGTCTCACCACCAAGTAACAAAAAGTAT 4320
Db 8260 CACATGATTTTCTTTTGTGTTTCTTCAATTTCTCTTGTCTCACCACCAAGTAACAAAAAGTAT 8319
QY 4321 AGTTTGAACAGAGTTGGTGTTCATAATTTCAATTTCACTTCTAGTTGATGCGAATTTTCAAT 4380
Db 8320 AGTTTGAACAGAGTTGGTGTTCATAATTTCAATTTCACTTCTAGTTGATGCGAATTTTCAAT 8379
QY 4381 AAGGAAGAGGGGCTTTTATCTTCTGCTAGGAAACCATGACGGAAGGAAACCTGAT 4440
Db 8380 AAGGAAGAGGGGCTTTTATCTTCTGCTAGGAAACCATGACGGAAGGAAACCTGAT 8439
QY 4441 GTTTAAAGTCCACATTTTAAACTATATTTATTTATGTTAGGATCTGTCAAAGAAAACCTTC 4500
Db 8440 GTTTAAAGTCCACATTTTAAACTATATTTATTTATGTTAGGATCTGTCAAAGAAAACCTTC 8499
QY 4501 CAAAAGATTTATTAATTAACACAGATCTGTGTGCAATTAAGTTAATGTTTCTTGTGTTT 4560
Db 8500 CAAAAGATTTATTAATTAACACAGATCTGTGTGCAATTAAGTTAATGTTTCTTGTGTTT 8559
QY 4561 TAATCCACACATTTCAATGAGTTAGGCTTTGCACTTTGTAAGGAAGGAGGCGGTTCACAC 4620
Db 8560 TAATCCACACATTTCAATGAGTTAGGCTTTGCACTTTGTAAGGAAGGAGGCGGTTCACAC 8619
QY 4621 CTCAAATAGCTAATAAACCGGCTCTTGAATATTTGAAGATTTTAAATCTCTGACTTAGGACG 4680
Db 8620 CTCAAATAGCTAATAAACCGGCTCTTGAATATTTGAAGATTTTAAATCTCTGACTTAGGACG 8679
QY 4681 GGCACGGTGGCTCAGCATATTAATCCCAACACTTTGGGAGGCTGAGGCGGGGTGCACAA 4740
Db 8680 GGCACGGTGGCTCAGCATATTAATCCCAACACTTTGGGAGGCTGAGGCGGGGTGCACAA 8739
QY 4741 GGTGAGGAGTTCAAGACCGACCTGACCAATATGTTGAACCCCATCTCTACTAAATAATAC 4800
Db 8740 GGTGAGGAGTTCAAGACCGACCTGACCAATATGTTGAACCCCATCTCTACTAAATAATAC 8799
QY 4801 AAAAAATTAGCCAGGCGTGGTGGAGGCTGCTGTAGTGTCCAGCTAGCCCTGTGAGGTGGAG 4860
Db 8800 AAAAAATTAGCCAGGCGTGGTGGAGGCTGCTGTAGTGTCCAGCTAGCCCTGTGAGGTGGAG 8859
QY 4861 ATTGCATTGAGCCAAAGATC 4879
Db 8860 ATTGCATTGAGCCAAAGATC 8878

RESULT 5
ADP65677
ID ADP65677 standard; DNA; 8878 BP.
XX AC ADP65677;
XX XX
DT 12-AUG-2004 (first entry)
XX XX
DE Human fibrinogen beta chain (FGB), complete cds DNA.
XX XX
KW autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; anti-rheumatic;
KW antiarthritic; osteopathic; antigout; anti-inflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; ds; human.
XX OS Homo sapiens.
XX XX
PN WO2003072827-A1.
XX XX
PD 04-SEP-2003.
XX XX
PF 31-OCT-2002; 2002WO-US035433.
XX XX
PR 31-OCT-2001; 2001US-0336220P.
XX XX

Qy	1501	GTGTGCATCTCGGTACTGGCTTTGAAACAGATGGCAGAACCCAGATGATGTAGAAGTTTC	1560
Db	5500	GTGTGCATCTCGGTACTGGCTTTGAAACAGATGGCAGAACCCAGATGATGTAGAAGTTTC	5559
Qy	1561	TCCATTTTGTGTGTTCTGGGAGCTCATGGAATATTCACAGACACAAAAGGTGGAGAGGC	1620
Db	5560	TCCATTTTGTGTGTTCTGGGAGCTCATGGAATATTCACAGACACAAAAGGTGGAGAGGC	5619
Qy	1621	TTTGTTCATCCTTTAGCAGATAAACCGTCTCAAAACTGGGTTTGGCACTTACTAAAGTAAA	1680
Db	5620	TTTGTTCATCCTTTAGCAGATAAACCGTCTCAAAACTGGGTTTGGCACTTACTAAAGTAAA	5679
Qy	1681	ATGAAATCTAAATATTTGTTATATATTTTCAAAGGCTATATATAACACACTCTCTTAGTA	1740
Db	5680	ATGAAATCTAAATATTTGTTATATATTTTCAAAGGCTATATATAACACACTCTCTTAGTA	5739
Qy	1741	ACTTATGTAATGTTATTTTAAAGAAATGGTGACTTAAATACAAAGTAATATGTCATAAAC	1800
Db	5740	ACTTATGTAATGTTATTTTAAAGAAATGGTGACTTAAATACAAAGTAATATGTCATAAAC	5799
Qy	1801	CCCTGGAACATAAATGTTGCTTTACATTTGCAGAATGTGAGAAATTAATACGAAAGGAGGT	1860
Db	5800	CCCTGGAACATAAATGTTGCTTTACATTTGCAGAATGTGAGAAATTAATACGAAAGGAGGT	5859
Qy	1861	GAAACATCTGAAATGTATCTCATTTCAACCTGCACAGTTTCTGTCAAACCGTATAGATATAC	1920
Db	5860	GAAACATCTGAAATGTATCTCATTTCAACCTGCACAGTTTCTGTCAAACCGTATAGATATAC	5919
Qy	1921	TGTGACATGAATACAGAAAAATGGAGGTAAAGCTTTTCGACAGTTGTGTGACCTGTTGATCTGT	1980
Db	5920	TGTGACATGAATACAGAAAAATGGAGGTAAAGCTTTTCGACAGTTGTGTGACCTGTTGATCTGT	5979
Qy	1981	AAATATTTTGGATACCGTAAAATGCCAGGAACAAAGCCAGGTGTGTGGCTCATACCTGT	2040
Db	5980	AAATATTTTGGATACCGTAAAATGCCAGGAACAAAGCCAGGTGTGTGGCTCATACCTGT	6039
Qy	2041	AAATTCACGACCTTTGGAGGCCAAAGTGGGCTGATAGCTTTGAGCCTTAGGAGTTTGAACCT	2100
Db	6040	AAATTCACGACCTTTGGAGGCCAAAGTGGGCTGATAGCTTTGAGCCTTAGGAGTTTGAACCT	6099
Qy	2101	AGCCTGGGCCAACATAATAGACACCTTAACCTCTACAAAAAATAAATAACAAAAAATAA	2160
Db	6100	AGCCTGGGCCAACATAATAGACACCTTAACCTCTACAAAAAATAAATAACAAAAAATAA	6159
Qy	2161	AAAAAATACAGCTGTGTGTTGGTAGTAGTGCTGTAGTCCAGCTATCCAGGAGGCTGAGA	2220
Db	6160	AAAAAATACAGCTGTGTGTTGGTAGTAGTGCTGTAGTCCAGCTATCCAGGAGGCTGAGA	6219
Qy	2221	TGGGAGATACCTTGAGCCCAACCTTGGAGCTTTGATCATGCTTACTGAACTGTAGCTTG	2280
Db	6220	TGGGAGATACCTTGAGCCCAACCTTGGAGCTTTGATCATGCTTACTGAACTGTAGCTTG	6279
Qy	2281	GCAACAGAGGATGTAGAGTCTCTGTCTCAAAAAAATAAATAAATAAAGCCAGGAAC	2340
Db	6280	GCAACAGAGGATGTAGAGTCTCTGTCTCAAAAAAATAAATAAATAAAGCCAGGAAC	6339
Qy	2341	AAGACTTAGCTCTAAACATCTAACATAGCTGCACAAAGGAGTAATTTGATGTGGAATTCAC	2400
Db	6340	AAGACTTAGCTCTAAACATCTAACATAGCTGCACAAAGGAGTAATTTGATGTGGAATTCAC	6399
Qy	2401	CTGATATTTTAAAGTTTATRAAATAATCTATAATTCACAAATTTGGGTAAAGTAAGCACTT	2460
Db	6400	CTGATATTTTAAAGTTTATRAAATAATCTATAATTCACAAATTTGGGTAAAGTAAGCACTT	6459
Qy	2461	GCAGTTTCCAAAGATTTTACAAGTTTACCTCTCATATTTTATTTTCCCTATTTTGTCTATTT	2520
Db	6460	GCAGTTTCCAAAGATTTTACAAGTTTACCTCTCATATTTTATTTTCCCTATTTTGTCTATTT	6519
Qy	2521	TAGAGCCACCAATATATACTAAATGGATGGACAGGGGATTCAGATATTTATTTTCAAAGT	2580
Db	6520	TAGAGCCACCAATATATACTAAATGGATGGACAGGGGATTCAGATATTTATTTTCAAAGT	6579

Qy	2581	GACATTA	TTTGC	TGTTGG	TAAATATATAG	CTCTTTTGGT	TCTGTCAACCAAGGATGGAC	2644
Db	6580	GACATTA	TTTGC	TGTTGG	TAAATATATAG	CTCTTTTGGT	CTGTCAACCAAGGATGGAC	6639
Qy	2641	AGTGATT	CAGAA	CGTCAAG	CGGTAG	TGTGACTTTGGC	GAGAAATGGGATCCATATAA	2700
Db	6640	AGTGATT	CAGAA	CGGTCAAG	CGGTAG	TGTGACTTTGGC	GAGAAATGGGATCCATATAA	6699
Qy	2701	ACAGGGA	TTTGG	AAATGTTCC	AAACCAAC	CAGATGGGA	GAATTA	2760
Db	6700	ACAGGGA	TTTGG	AAATGTTCC	AAACCAAC	CAGATGGGA	GAATTA	6759
Qy	2761	TAACGA	ACAGC	ATGCAAA	TAAATCA	TTCTTA	TGAAATGGATTTTTTTTAA	2820
Db	6760	TAACGA	ACAGC	ATGCAAA	TAAATCA	TTCTTA	TGAAATGGATTTTTTTTAA	6819
Qy	2821	AAACATT	CATTG	TTGGA	AGCCTGTTT	TAGCAG	TAAAGAGGATTTCTC	2880
Db	6820	AAACATT	CATTG	TTGGA	AGCCTGTTT	TAGCAG	TAAAGAGGATTTCTC	6879
Qy	2881	GGAAGCT	TAAGAT	TAAGGA	AAAGGCAG	TTTTAG	TTTCCCAAAATTTATTTTGGTG	2940
Db	6880	GGAAGCT	TAAGAT	TAAGGA	AAAGGCAG	TTTTAG	TTTCCCAAAATTTATTTTGGTG	6939
Qy	2941	AGAGATT	TTATTTG	TTTTCTTTT	TAGGTGAA	TATTTGGCT	TGAAATGATAAAATATAGCC	3000
Db	6940	AGAGATT	TTATTTG	TTTTCTTTT	TAGGTGAA	TATTTGGCT	TGAAATGATAAAATATAGCC	6999
Qy	3001	AGCTTAC	CGGATGGGA	CCACAC	AGAACTTTT	TGATAGAA	ATGGAAGGAGAC	3060
Db	7000	AGCTTAC	CGGATGGGA	CCACAC	AGAACTTTT	TGATAGAA	ATGGAAGGAGAC	7059
Qy	3061	AAGTAA	AGGCTCA	CTATG	GAGGATTC	ACTCTAC	AGAAATGAAGCCAAATATCCAGATCT	3120
Db	7060	AAGTAA	AGGCTCA	CTATG	GAGGATTC	ACTCTAC	AGAAATGAAGCCAAATATCCAGATCT	7119
Qy	3121	CAGTGAA	CAAAATAC	AGAGGA	ACCGCGTAA	TCCCTCAT	TGATGGAGCATCTCAGCTGA	3180
Db	7120	CAGTGAA	CAAAATAC	AGAGGA	ACCGCGTAA	TCCCTCAT	TGATGGAGCATCTCAGCTGA	7179
Qy	3181	TGGGAG	AAACCGG	ACCATG	ACCACTTCA	CAACGGCATG	TTCTCAGACGTATGACAG	3240
Db	7180	TGGGAG	AAACCGG	ACCATG	ACCACTTCA	CAACGGCATG	TTCTCAGACGTATGACAG	7239
Qy	3241	ACAATGA	CGGCTGG	TATGTGTG	GCAC	CTTTGCTCCT	TGCTTTAAAAATCACACTAATATC	3300
Db	7240	ACAATGA	CGGCTGG	TATGTGTG	GCAC	CTTTGCTCCT	TGCTTTAAAAATCACACTAATATC	7299
Qy	3301	ATTACT	CAGAAATCA	TTAA	CAATATTTT	TAATAGCTA	CCACTTCTCGGGCACTTACTGTCA	3360
Db	7300	ATTACT	CAGAAATCA	TTAA	CAATATTTT	TAATAGCTA	CCACTTCTCGGGCACTTACTGTCA	7359
Qy	3361	GCCACT	GTCTTA	AGCTCTTT	TATGCATC	ACCTCGAAAGCATTT	CAACTATAAGGTAGACATT	3420
Db	7360	GCCACT	GTCTTA	AGCTCTTT	TATGCATC	ACCTCGAAAGCATTT	CAACTATAAGGTAGACATT	7419
Qy	3421	CTTATTC	TCAATTTT	CAGATG	AGATTTAG	AGAGATTA	CCGTGATTTGCCAATGTCAACA	3480
Db	7420	CTTATTC	TCAATTTT	CAGATG	AGATTTAG	AGAGATTA	CCGTGATTTGCCAATGTCAACA	7479
Qy	3481	ACTAC	CCAGAGATAA	ACCTAG	ATTTGAG	CAACAGTTT	CTTCTGAAATATGACATTTAG	3540
Db	7480	ACTAC	CCAGAGATAA	ACCTAG	ATTTGAG	CAACAGTTT	CTTCTGAAATATGACATTTAG	7539
Qy	3541	ATAAAT	ACCTATATCT	CTATATTT	CTAAAG	TGTGTG	TGAAAACTTTTCATTTTCCAG	3600
Db	7540	ATAAAT	ACCTATATCT	CTATATTT	CTAAAG	TGTGTG	TGAAAACTTTTCATTTTCCAG	7599
Qy	3601	GGTCTC	TGATCTAG	GGTTGT	TAAAGCTAT	TATTTCC	AGTATAAGTAAACACAGT	3660
Db	7600	GGTCTC	TGATCTAG	GGTTGT	TAAAGCTAT	TATTTCC	AGTATAAGTAAACACAGT	7659
Qy	3661	CCCTAG	TGATTTGCC	CAAAAGGCC	ACGTTATCT	CTCTTTT	CTTGCTATAGGGCACAGGAG	3720

CC associated with abnormalities in the expression, amount or distribution
CC of the protein. The present sequence represents a cDNA of the invention
CC that is differentially expressed in activated vascular tissue. Note: The
CC sequence data for this patent did not form part of the specification, but
CC was obtained in electronic format directly from USPTO at
CC <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>

Sequence	8674 BP; 2658 A; 1311 C; 1485 G; 2640 T; 0 U; 580 Other;
Query Match	87.5%; Score 4268.2; DB 8; Length 8674;
Best Local Similarity	92.7%; Pred. No. 0;
Matches 4336; Conservative	0; Mismatches 336; Indels 6; Gaps 6;
Qy	1 TCAGAAATAGTGTGATGAGTTAAATACAAATGTGGAAGCTGTTTCCAGACCTCTCTTT 60
Db	4000 TCAGAAATAGTGTGATGAGTTAAATACAAATGTGGAAGCTGTTTCCAGACCTCTCTTT 4059
Qy	61 CTTTCCTTTTCAGTACATGATTTTGCTGAAAGACCTGTGCAAAAGAGCAGACGACGAAGTAA 120
Db	4060 CTTTCCTTTTCAGTACATGATTTTGCTGAAAGACCTGTGCAAAAGAGCAGACGAGTAA 4119
Qy	121 AAGGTAGATATCCTTTGTGCTTTCCATTCGATTTTCAGTATAAATTTGGAAACCGTTAGAC 180
Db	4120 AAGGTAGATATCCTTTGTGCTTTCCATTCGATTTTCAGTATAAATTTGGAAACCGTTAGAC 4179
Qy	181 TGGCACAGAAATGCATGGTTGTGAGAGATTAAATTTCTGGGTTAGTGAATAGCATTC A 240
Db	4180 TGGCACAGAAATGCATGGTTGTGAGAGATTAAATTTCTGGGTTAGTGAATAGCATTC A 4239
Qy	241 TACGCTTTTGGGCACCTTCCCTCGCAATTCGCCAGATAAGCACATTTTCAGCTCTTATTC 300
Db	4240 TACGCTTTTGGGCACCTTCCCTCGCAATTCGCCAGATAAGCACATTTTCAGCTCTTATTC 4299
Qy	301 CAGTCTGCATCAGCAAGTGTGATTTTCTATGAAAAATTTCTACTATGACTCCTTTATTTTA 360
Db	4300 CAGTCTGCATCAGCAAGTGTGATTTTCTATGAAAAATTTCTACTATGACTCCTTTATTTTA 4359
Qy	361 AGTATACAGAAACTGTGTGACTCAGAGAGATAATTTTACAGAGTGGAAAAAACCCTTAG 420
Db	4360 AGTATACAGAAACTGTGTGACTCAGAGAGATAATTTTACAGAGTGGAAAAAACCCTTAG 4419
Qy	421 CATTTATAGTTTTAAACATTTTGAGGTTTTGAATGAGAGAGTTTATCCATAATATTTCAATT 480
Db	4420 CATTTATAGTTTTAAACATTTTGAGGTTTTGAATGAGAGAGTTTATCCATAATATTTCAATT 4479
Qy	481 GTGTTGTGGATAATGCACACCTTAACCTGTGATCTTTGAGGTGAGAAATGTTGAGTCTGTTG 540
Db	4480 GTGTTGTGGATAATGCACACCTTAACCTGTGATCTTTGAGGTGAGAAATGTTGAGTCTGTTG 4539
Qy	541 ACTTTGGTGTGAGAAACAGCTAGTGGTGGAGCCTGGCACAGGCATCTCAGTGAAGTACGA 600
Db	4540 ACTTTGGTGTGAGAAACAGCTAGTGGTGGAGCCTGGCACAGGCATCTCAGTGAAGTACGA 4599
Qy	601 TACCACAGTTGGAAAAATTTTTCAAGAGAAATCAAAGGAATCATGACATCTCTTATAAATTTCA 660
Db	4600 TACCACAGTTGGAAAAATTTTTCAAGAGAAATCAAAGGAATCATGACATCTCTTATAAATTTCA 4659
Qy	661 AGGTTCTGCTATACTTATGTGAAATGGATAAATAAATCAAGCATATCCACTCTGTAAGAT 720
Db	4660 AGGTTCTGCTATACTTATGTGAAATGGATAAATAAATCAAGCATATCCACTCTGTAAGAT 4719
Qy	721 TGAATCTTCAGATGGAGACCCCAATACCTGCTTTCTCCTCTTTTCCCTCACCAGAGAAA 780
Db	4720 TGAATCTTCAGATGGAGACCCCAATACCTGCTTTCTCCTCTTTTCCCTCACCAGAGAAA 4779
Qy	781 TAAACAACTATTTCAATTTATTTACTGGACACAACTTTTAGCGTATACCTATGTAATTA 840
Db	4780 TAAACAACTATTTCAATTTATTTACTGGACACAACTTTTAGCGTATACCTATGTAATTA 4839
Qy	841 CTAGTATCGGTGTAGGAATTTATGTTAATTTGTATATGTGATGGCGCAAAATCATTTCCAC 900
Db	4840 CTAGTATCGGTGTAGGAATTTATGTTAATTTGTATATGTGATGGCGCAAAATCATTTCCAC 4899

QY 4198 ACTCTCAAGCAGACGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAATACATTAAAA 4257
 Db 8198 ACTCTCAAGCAGACGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAATACATTAAAA 8257
 QY 4258 TAGCACATGATTTCTTTTGTGTTCTTCAATTTCTCTGCTACCCAGAAAGTAAACAAAG 4317
 Db 8258 TAGCACATGATTTCTTTTGTGTTCTTCAATTTCTCTGCTACCCAGAAAGTAAACAAAG 8316
 QY 4318 TATAGTTTTCAGACAGTGTGCTTCATTAATTTCACTTCTAGTTGATTCGAGAAATTTCA 4377
 Db 8317 TATAGTTTTCAGACAGTGTGCTTCATTAATTTCACTTCTAGTTGATTCGAGAAATTTCA 8376
 QY 4378 AATAAGGAAGGGGCTTTTATCTCTGCTAGGAAACCATGACGGAAGGAAAGAACT 4437
 Db 8377 AATAAGGAAGGGGCTTTTATCTCTGCTAGGAAACCATGACGGAAGGAAAGAACT 8436
 QY 4438 GATGTTTAAAGTCCACATTTTAAACTATATTTATTTAGTATGATGATCTGTCAAGAAAC 4497
 Db 8437 GATGTTTAAAGTCCACATTTTAAACTATATTTATTTATGATGATCTGTCAAGAAAC 8496
 QY 4498 TTCCAAAAGATTTTAAATTAACACGACTCTGTTGCAATAAGTTAATGTTTCTTCTT 4557
 Db 8497 TTCCAAAAGATTTTAAATTAACACGACTCTGTTGCAATAAGTTAATGTTTCTTCTT 8556
 QY 4558 TTGTAATCCACACATTCATGAGTTAGGCTTTGCACTTTGTAAGGAAGGAGCGTTTCA 4617
 Db 8557 TTGTAATCCACACATTCATGAGTTAGGCTTTGCACTTTGTAAGGAAGGAGCGTTTCA 8616
 QY 4618 AACCTCAATAGCTAATAACCGGCTTGAATATTTGAAGATTTAAATCTGACTCTA 4675
 Db 8617 AACCTCAATAGCTAATAACCGGCTTGAATATTTGAAGATTTAAATCTGACTCTA 8674

RESULT 7

ABX91986
 ID ABX91986 standard; cDNA; 4043 BP.

XX AC ABX91986;

XX DT 08-MAY-2003 (first entry)

XX DE Lung specific nucleic acid (LSNA) #28.

XX XX Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;
 KW cancer monitoring; cancer staging; cancer imaging; lung cancer;
 KW non-cancerous diseases of the lung; transgenic animal; gene; ss.

XX OS Homo sapiens.

XX PN WO200268633-A2.

XX PD 06-SEP-2002.

XX PF 21-NOV-2001; 2001WO-US043612.

XX PR 22-NOV-2000; 2000US-0252500P.

XX XX (DIAD-) DIADEXUS INC.

XX PI Macina RA, Recipon H, Chen S, Sun Y, Liu C;

XX DR WPI; 2002-713376/77.

XX XX New isolated human nucleic acid molecule and polypeptide, useful for
 PT identifying, diagnosing, monitoring, staging, imaging and treating lung
 PT cancer and non-cancerous diseases of the lung.

XX PS Claim 1; Page 181-183; 389pp; English.

XX XX The invention describes an isolated human nucleic acid (I) encoding any
 CC of 120-1533 residue amino acid sequences (SI), given in the
 CC specification, comprising any of 164 179-12421 base pair sequences (S2),

CC given in the specification. The methods and compositions of the present
 CC invention are useful for identifying, diagnosing, monitoring, staging,
 CC imaging and treating lung cancer and non-cancerous diseases of the lung.
 CC They are also used for identifying lung tissue, monitoring and
 CC identifying and/or designing antagonists of the polypeptide of the
 CC invention, gene therapy, production of transgenic animals and production
 CC of engineered lung tissue for treatment and research. This sequence
 CC encodes a lung specific nucleic acid

XX SQ Sequence 4043 BP; 1283 A; 824 C; 842 G; 1094 T; 0 U; 0 Other;

Query Match 42.7%; Score 2081.2; DB 6; Length 4043;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2083; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2018 CAGGTGTGGTGGCTCATACCTGTAATCCACAGCCTTGGGAGGCCAAAGTGGCTGTAG 2077
 Db 1958 CAAGTGTGGTGGCTCACACCTGTAATCCACAGCCTTGGGAGGCCAAAGTGGCTGTAG 2017
 QY 2078 CTTGAGCCTTAGGAGTTTGAAACTAGCCTGGGCAACATATGAGACCCCTAACTTACAAA 2137
 Db 2018 CTTGAGCCTTAGGAGTTTGAAACTAGCCTGGGCAACATATGAGACCCCTAACTTACAAA 2077
 QY 2138 AAAAAAATPACAAAAAATAAATCAGCTGTGTGGTAGTATGTCCTGTAGT 2197
 Db 2078 AAAAAAATPACAAAAAATAAATCAGCTGTGTGGTAGTATGTCCTGTAGT 2137
 QY 2198 CCCAGCTATCCAGAGGCTGAGATGGGAGATCCTGAGCCCAACCTGGAGTCTCAT 2257
 Db 2138 CCCAGCTATCCAGAGGCTGAGATGGGAGATCCTGAGCCCAACCTGGAGTCTCAT 2197
 QY 2258 CATGCTACTGAACCTGTAGCCTGGGCAACAGAGGATAGTGAGATCCTGCTCAAAAAAAA 2317
 Db 2198 CATGCTACTGAACCTGTAGCCTGGGCAACAGAGGATAGTGAGATCCTGCTCAAAAAAAA 2257
 QY 2318 AATTAATTAAGCCAGGAAACAAAGACTTAGCTCTTAACATCTTAACATAGCTGACAAAG 2377
 Db 2258 AATTAATTAAGCCAGGAAACAAAGACTTAGCTCTTAACATCTTAACATAGCTGACAAAG 2317
 QY 2378 AGTAATTTGATGTGGAAATTCACCTGATATTTAAAGTTATAAATATCTAATTTACA 2437
 Db 2318 AGTAATTTGATGTGGAAATTCACCTGATATTTAAAGTTATAAATATCTAATTTACA 2377
 QY 2438 ATTTGGGTAAGATAAAGCACTTGCAGTTTCCAAAGATTTTCAAGTTTACCTCTCATAT 2497
 Db 2378 ATTTGGGTAAGATAAAGCACTTGCAGTTTCCAAAGATTTTCAAGTTTACCTCTCATAT 2437
 QY 2498 TTATTTCTTATTTGTCTATTTTAGAGCAACAAATATATATCTAAATGGAAATGGACAGG 2557
 Db 2438 TTATTTCTTATTTGTCTATTTTAGAGCAACAAATATATATCTAAATGGAAATGGACAGG 2497
 QY 2558 GATTCAGATATTTTTCAAAGTGCATTTATTTGCTGTGGTTAATATATCTCTTTTGG 2617
 Db 2498 GATTCAGATATTTTTCAAAGTGCATTTATTTGCTGTGGTTAATATATCTCTTTTGG 2557
 QY 2618 TTTCTGTCAACCAAGGATGGACAGTGAATTCAGAACCGCTCAAGCGTAGTGTGACTTT 2677
 Db 2558 TTTCTGTCAACCAAGGATGGACAGTGAATTCAGAACCGCTCAAGCGTAGTGTGACTTT 2617
 QY 2678 GGCAGGAAATGGGATCCATATAAAGCAGGATTTGGAAATTTGGCAACCAACACAGATGGG 2737
 Db 2618 GGCAGGAAATGGGATCCATATAAAGCAGGATTTGGAAATTTGGCAACCAACACAGATGGG 2677
 QY 2738 AAGAAATTTACTGTGGCTTACCAAGTAAAGCAAGCAGCATGCAAAATATAATCTATTATTG 2797
 Db 2678 AAGAAATTTACTGTGGCTTACCAAGTAAAGCAAGCAGCATGCAAAATATAATCTATTATTG 2737
 QY 2798 AAATGGGATTTTTTTTAAATTAATAAACAATTTGTTGGAGCCCTGTTTTAGGCAGTTAA 2857
 Db 2738 AAATGGGATTTTTTTTAAATTAATAAACAATTTGTTGGAGCCCTGTTTTAGGCAGTTAA 2797
 QY 2858 GAGAGTTTCTGTGACAAAAATGCTGAGAGCTTAAGATTAAGGAAGAAAGCGAGTTTTT 2917

2978 GAGGAGTTTCTGACAAAAATGTGGAGCTAAAGATAAGGGAAGAAAGGAGCTTTTACT 2857
2918 TTCCCAAAATTTATTTTGGTGGAGATTTATTTGTTTCTTTTATAGTGAATATTG 2977
2858 TTCCCAAAATTTATTTTGGTGGAGATTTATTTGTTTCTTTTATAGTGAATATTG 2917
2978 GCTTGGAAATGATAAAATTTAGCCAGCTTACCAGATGGGACCCACAGAACTTTTGTATAGA 3037
2918 GCTTGGAAATGATAAAATTTAGCCAGCTTACCAGATGGGACCCACAGAACTTTTGTATAGA 2977
3038 AATGGAGGCTGAAAGGAGACAAAGTAAAGGCTCACTATGGAGGATTCAGTGTACAGAA 3097
2978 AATGGAGGCTGAAAGGAGACAAAGTAAAGGCTCACTATGGAGGATTCAGTGTACAGAA 3037
3098 TGAAGCCAAATATACCATCTCTCAGTGAACAAATACAGGAAACGCGGTAATGCCCT 3157
3038 TGAAGCCAAATATACCATCTCTCAGTGAACAAATACAGGAAACGCGGTAATGCCCT 3097
3158 CATGATGGAGCATCTCAGCTGATGGGAGAAAACAGGACCATGACCATTCACAACGGCAT 3217
3098 CATGATGGAGCATCTCAGCTGATGGGAGAAAACAGGACCATGACCATTCACAACGGCAT 3157
3218 GTTCTTCAGCAGTATGACAGAGACAATAGCGGCTGGTATGTTGGCACTCTTTGCTCCT 3277
3158 GTTCTTCAGCAGTATGACAGAGACAATAGCGGCTGGTATGTTGGCACTCTTTGCTCCT 3217
3278 GCTTTAAATATCACATATATCATCTACTCAGATCATTAACAATATTTTAAATAGCTAC 3337
3218 GCTTTAAATATCACATATATCATCTACTCAGATCATTAACAATATTTTAAATAGCTAC 3277
3338 CACTTCTGGGCACTTACTGTGAGCCACTGTCTTAAGCTCTTTATGCATCACTCGAAAGC 3397
3278 CACTTCTGGGCACTTACTGTGAGCCACTGTCTTAAGCTCTTTATGCATCACTCGAAAGC 3337
3398 ATTTCACATATAGGTAGACATCTTATTTCTCATTTTACAGATGAGATTTAGAGATTA 3457
3338 ATTTCACATATAGGTAGACATCTTATTTCTCATTTTACAGATGAGATTTAGAGATTA 3397
3458 CBTGATTTGTCGAATGTACACAACTACCCAGAGATAAACTAGATTTGACACAGTTA 3517
3398 CBTGATTTGTCGAATGTACACAACTACCCAGAGATAAACTAGATTTGACACAGTTA 3457
3518 CTTTCTGAATATGACATTTAGATAAATACCTATATCTCTATATTTCTAAAGTGTGTGG 3577
3458 CTTTCTGAATATGACATTTAGATAAATACCTATATCTCTATATTTCTAAAGTGTGTGG 3517
3578 AAAAATTTTCATTTTCCAGGGTCTCTGATACCTAAGGGTGTGTAAGGCTATTTATTC 3637
3518 AAAAATTTTCATTTTCCAGGGTCTCTGATACCTAAGGGTGTGTAAGGCTATTTATTC 3577
3638 CAGTATAAAGTAAACACACAGTCCCTAGATGATTCACAAAGGCCAGTATCTCTC 3697
3578 CAGTATAAAGTAAACACACAGTCCCTAGATGATTCACAAAGGCCAGTATCTCTC 3637
3698 TTTCTTGCTATAGGACACAGGAGTCTTTGGTGTATTTAGTGTACCTATGTATAGCAC 3757
3638 TTTCTTGCTATAGGACACAGGAGTCTTTGGTGTATTTAGTGTACCTATGTATAGCAC 3697
3758 CAAAGAAAGACTACTGTGACACAGAGTAGCAGTCTTTTATGGTAAATCTGCAAAAGC 3817
3698 CAAAGAAAGACTACTGTGACACAGAGTAGCAGTCTTTTATGGTAAATCTGCAAAAGC 3757
3818 TAACTTGACCCGTTAGTTCTGTTCTTATTAACGCCAAACACATTTTCTTCAGGTAAAC 3877
3758 TAACTTGACCCGTTAGTTCTGTTCTTATTAACGCCAAACACATTTTCTTCAGGTAAAC 3817
3878 ATCAGATCCAGAAAAACAGTGTCTTAAAGAACGCGTGTGGATGTGTATTAATAGATG 3937
3818 ATCAGATCCAGAAAAACAGTGTCTTAAAGAACGCGTGTGGATGTGTATTAATAGATG 3877
3938 TCATGACGCAATCCAAACGGCAGATACCTACTGGGTGGACAGTACACCTGGGACATGGC 3997
3878 TCATGACGCAATCCAAACGGCAGATACCTACTGGGTGGACAGTACACCTGGGACATGGC 3937

QY 3998 AAGCATGCGCACAGATGATGGTGTAGTATGATGAATTCGAAGGGTCACTGGTACTCAAT 4057
Db 3938 AAGCATGCGCACAGATGATGGTGTAGTATGATGAATTCGAAGGGTCACTGGTACTCAAT 3997
QY 4058 GAGGAAGATGAGTATGATGAAGATCAGGCCCTTCTTCCACAGCAATAG 4103
Db 3998 GAGGAAGATGAGTATGATGAAGATCAGGCCCTTCTTCCACAGCAATAG 4043

RESULT 8

AAS60957/C

ID AAS60957 standard; cDNA; 4967 BP.

XX AAS60957;

AC AAS60957;

XX 29-JAN-2002 (first entry)

XX Human cancer agent-resistance marker #616.

XX Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;

XX squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;

XX lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;

XX Hodgkin's disease; glioma; ss.

XX Homo sapiens.

XX WO200179556-A2.

XX 25-OCT-2001.

XX 13-APR-2001; 2001WO-US012132.

XX 14-APR-2000; 2000US-0197538P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Brown JL, Bolt A, Van Huffel C;

XX WPI; 2001-602933/68.

XX Novel nucleic acid, used as a marker to determine the effectiveness of

XX using TAXOL to treat cancer cell growth in individuals.

XX Claim 1; Page 459-460; 527pp; English.

XX The invention relates to 1046 novel nucleic acids which are used as

XX markers for determining the sensitivity of a cancer cell to the

XX anticancer agent TAXOL. Cancer cells can be treated with TAXOL when they

XX are shown to express one of the 242 sensitivity markers of the cells are

XX shown not to express one of the 804 resistance markers. The methods can

XX be used to determine the effectiveness of TAXOL in the treatment of

XX cancer cell growth in an individual. The markers can be used as targets

XX in developing anti-cancer agents such as chemotherapeutic compounds. The

XX markers can also be used as targets in developing treatments for cancer.

XX particularly those cancers which display resistance to agents and exhibit

XX expression of the markers. The anticancer agents developed by the novel

XX method can be used to treat cancer. Probes based on the markers can be

XX used to detect transcripts or genomic sequences corresponding to the

XX markers, in the identification of cells or tissues which mis-express the

XX protein. Cancers which may be targeted include carcinoma (e.g. squamous

XX cell carcinoma), sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic

XX leukaemia), lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's

XX disease and tumours (e.g. glioma). The present sequence is one of the

XX 1046 novel cancer cell markers

XX Sequence 4967 BP; 1232 A; 1163 C; 1177 G; 1395 T; 0 U; 0 Other;

XX Query Match 19.9%; Score 971.8; DB 4; Length 4967;

XX Best Local Similarity 99.5%; Pred. No. 9.8e-194;

XX Matches 1006; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 3869 CAGGTAAATCATCAGATCCAGAAACAGTGTCTTAAAGAACGCGTGTGGATGGTGA 3928

Db 2428 CTGTTAAACATCAGATCCAGAAAAACAGTTCTAAAGAACGGTGTGGTGGTA 2369
Qy 3929 TAATAGATGTCATGAGCAATCCAAACGGCAGATCTACTGGGTGGACAGTACACCTG 3988
Db 2368 TAATAGATGTCATGAGCAATCCAAACGGCAGATCTACTGGGTGGACAGTACACCTG 2309
Qy 3989 GGACATGGCAACAGCATGGCACAGATGATGGTGTAGTATGATGAATTTGGAAGGGTCAATG 4048
Db 2308 GGACATGGCAACAGCATGGCACAGATGATGGTGTAGTATGATGAATTTGGAAGGGTCAATG 2249
Qy 4049 GTACTCAATAGGAAGATGAGTATGAAGATCAGGCGCTTCTCCACAGCAATAGTCCCC 4108
Db 2248 GTACTCAATAGGAAGATGAGTATGAAGATCAGGCGCTTCTCCACAGCAATAGTCCCC 2189
Qy 4109 AATAGTATAGTATTTCTCTCTGTATGTGACAAACATTTTGTACATATGTTATTTGGA 4168
Db 2188 AATAGTATAGTATTTCTCTCTGTATGTGACAAACATTTTGTACATATGTTATTTGGA 2129
Qy 4169 TTTTCTTTTATATATATTTCTCTCTAAACTCTCAAGCAGACGTGAGTGTGACTTTTGTG 4228
Db 2128 TTTTCTTTTATATATATTTCTCTCTAAACTCTCAAGCAGACGTGAGTGTGACTTTTGTG 2069
Qy 4229 AAAAAAGTATAGGATAAATTAATTAATTAATAGCAGATGATTTTCTTTTCTTTCTCAT 4288
Db 2068 AAAAAAGTATAGGATAAATTAATTAATTAATAGCAGATGATTTTCTTTTCTTTCTCAT 2009
Qy 4289 TCTCTTGTCTACCCAGGAAGTAAACAGTATAGTTTTCACAGAGTGGTGTTCATAAT 4348
Db 2008 TCTCTTGTCTACCCAGGAAGTAAACAGTATAGTTTTCACAGAGTGGTGTTCATAAT 1950
Qy 4349 TCAGTTCTAGTTGATTTGCGAGAAATTTTCAAAAGAGAGGGGTCTTTTATCCTTTGTCG 4408
Db 1949 TCAGTTCTAGTTGATTTGCGAGAAATTTTCAAAAGAGAGGGGTCTTTTATCCTTTGTCG 1890
Qy 4409 TAGGAAACCATGACGGAAGGAAGAACTGATGTTTAAAGTCCACTTTTAAACATATAT 4468
Db 1889 TAGGAAACCATGACGGAAGGAAGAACTGATGTTTAAAGTCCACTTTTAAACATATAT 1830
Qy 4469 TTATTTATGATGATCTGTCAGGAAGAACTTCCAAAGAGATTTATTAATTAACACGACT 4528
Db 1829 TTATTTATGATGATCTGTCAGGAAGAACTTCCAAAGAGATTTATTAATTAACACGACT 1770
Qy 4529 CTGTTGCAATAGTTAAATGTTTCTTTGTTTGTGTAATCCACATCAATGAGTTAGGCTT 4588
Db 1769 CTGTTGCAATAGTTAAATGTTTCTTTGTTTGTGTAATCCACATCAATGAGTTAGGCTT 1710
Qy 4589 TGCACCTTGTAGGAGGAGGAGGTTCAAACTCAATAGCTAATAAACCGGCTTGAA 4648
Db 1709 TGCACCTTGTAGGAGGAGGAGGTTCAAACTCAATAGCTAATAAACCGGCTTGAA 1650
Qy 4649 TATTTGAAGATTTAAATCTGACTCTAGGACGGCAGCGTGGCTCACGACTATATATCCCA 4708
Db 1649 TATTTGAAGATTTAAATCTGACTCTAGGACGGCAGCGTGGCTCACGACTATATATCCCA 1590
Qy 4709 ACATTTGGAGGCTGAGCGCGGCGGTCAAGAGTCAGAGTTCAAGACAGCCTGACCA 4768
Db 1589 ACATTTGGAGGCTGAGCGCGGCGGTCAAGAGTCAGAGTTCAAGACAGCCTGACCA 1530
Qy 4769 ATATGTTGAACCCCATCTCTACTAATAAATACAAATTTAGCCAGCGGTGGGAGGTTG 4828
Db 1529 ATATGTTGAACCCCATCTCTACTAATAAATACAAATTTAGCCAGCGGTGGGAGGTTG 1470
Qy 4829 CTTGTAGGTCACAGCTAGCTGTGAGGTGGAGATTCATTTAGCCAGCAAGATC 4879
Db 1469 CTTGTGTA-GTCCACAGCTA-CTGTGAGGTGGAGATTCATTTAGCCAGCAAGATC 1421

RESULT 9
ADL63650/c
ID ADL63650 standard; DNA; 4967 BP.
XX
AC ADL63650;

XX 20-MAY-2004 (first entry)
XX Human ovarian cancer DNA marker #21862.
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX Homo sapiens.
XX WO2001170979-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009126.
XX 21-MAR-2000; 2000US-0191031P.
XX 25-MAY-2000; 2000US-0207124P.
XX 15-JUN-2000; 2000US-0211940P.
XX 07-JUL-2000; 2000US-0216820P.
XX 25-JUL-2000; 2000US-0220661P.
XX 21-DEC-2000; 2000US-0257672P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Lee J, Lillie J;
XX WPI; 2001-611502/70.
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.
XX Disclosure; SEQ ID NO 21862; 106pp; English.
XX The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention.
XX Sequence 4967 BP; 1232 A; 1163 C; 1177 G; 1395 T; 0 U; 0 Other;
Qy 3869 CAGGTTAATCAGATCCAGAAACAGTTGTTCTAAAGAGACGGTGTGGATGGTGA 3928

Query Match 19.9%; Score 971.8; DB 5; Length 4967;
Best Local Similarity 99.5%; Pred. No. 9.8e-194;
Matches 1006; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
Qy 3869 CAGGTTAATCAGATCCAGAAACAGTTGTTCTAAAGAGACGGTGTGGATGGTGA 3928

QY 4529 CTGTTGCAATAAGTTAATGTTTCTT 4554
 Db |||||
 1915 CTTGTCATAAATAATATAGTTTT 1940
 |||||
 RESULT 12
 AAK52610
 ID AAK52610 standard; cDNA; 2553 BP.
 XX
 AC AAK52610;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 2139.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 03-FEB-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR P-PSDB; AM79477.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 1; Page 4509-4510; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 2553 BP; 812 A; 506 C; 546 G; 689 T; 0 U; 0 Other;
 Query Match 13.5%; Score 659.6; DB 4; Length 2553;
 Best Local Similarity 98.5%; Pred. No. 2.6e-128;
 Matches 676; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
 QY 3869 CAGGTTAATCATCAGATCCAGAAAACAGTGTCTTAAAGAACGCGTGGTGGTGGTA 3928

Db |||||
 1256 CTGTTTAAATCAGATCCAGAAAACAGTGTCTTAAAGAACGCGTGGTGGTGGTA 1315
 QY |||||
 3929 TAAATAGATGTCAATGACGCAATCCAAACGGCAGATACATCTAGGGGTGGACAGTACACCTG 3988
 Db |||||
 1316 TAAATAGATGTCAATGACGCAATCCAAACGGCAGATACATCTAGGGGTGGACAGTACACCTG 1375
 QY |||||
 3989 GGACATGGCAAGCATGGCAGATGATGGTGTAGTATGGATGAATGGAAAGGGGTGATG 4048
 Db |||||
 1376 GGACATGGCAAGCATGGCAGATGATGGTGTAGTATGGATGAATGGAAAGGGGTGATG 1435
 QY |||||
 4049 GTACTCAATGAGGAAGATGAGTATGAAGATCAGAGCCCTTCTCCACAGCAATAGTCCCC 4108
 Db |||||
 1436 GTACTCAATGAGGAAGATGAGTATGAAGATCAGAGCCCTTCTCCACAGCAATAGTCCCC 1495
 QY |||||
 4109 AATAGTATGATTTTGTCTCTCTATGTGACAAATTTTGTACATTAATGTTATGGAA 4168
 Db |||||
 1496 AATAGTATGATTTTGTCTCTCTATGTGACAAATTTTGTACATTAATGTTATGGAA 1555
 QY |||||
 4169 TTTTCTTTCATACATTAATTTCTCTTAAACTCTCAAGCAGACGTGAGTGTGACTTTTG 4228
 Db |||||
 1556 TTTTCTTTCATACATTAATTTCTCTTAAACTCTCAAGCAGACGTGAGTGTGACTTTTG 1615
 QY |||||
 4229 AAAAAATAGTATGATAAATTAATTAATAGCACAATGATTTCTTTTGTCTTCTTCAAT 4288
 Db |||||
 1616 AAAAAATAGTATGATAAATTAATTAATAGCACAATGATTTCTTTTGTCTTCTTCAAT 1675
 QY |||||
 4289 TCTTGTCTCACCAAGAGTAACAAAAGTATAGTGTGACAGAGTGGTGTGCTAAT 4348
 Db |||||
 1676 TCTTGTCTCA-CCAAAGAGTAACAAAAGTATAGTGTGACAGAGTGGTGTGCTAAT 1734
 QY |||||
 4349 TCAGTCTTAGTGTGATGGAGAAATTTCAATTAAGGAGAGGGGTCTTTATCCTGTG 4408
 Db |||||
 1735 TCAGTCTTAGTGTGATGGAGAAATTTCAATTAAGGAGAGGGGTCTTTATCCTGTG 1794
 QY |||||
 4409 TAGGAAAACCATGACGGAAGGAAAACCTGATGTTTAAAGTCCACTTTTAAACATATAT 4468
 Db |||||
 1795 TAGGAAAACCATGACGGAAGGAAAACCTGATGTTTAAAGTCCACTTTTAAACATATAT 1854
 QY |||||
 4469 TTATTTATGTAGGATCTGTCAAAGAAAACCTTCCAAAAGATTTATTAATAAACAGACT 4528
 Db |||||
 1855 TTATTTATGTAGGATCTGTCAAAGAAAACCTTCCAAAAGATTTATTAATAAACAGACT 1914
 QY |||||
 4529 CTGTTGCAATAGTTAATGTTTCTT 4554
 Db |||||
 1915 CTGTTGCAATAAATAATATAGTTTT 1940
 RESULT 13
 AAK51626
 ID AAK51626 standard; cDNA; 1862 BP.
 XX
 AC AAK51626;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 171.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR

PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
DR WPI; 2001-476283/51.
DR P-PSDB; AAM78493.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 1; Page 909-911; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
XX SQ Sequence 1862 BP; 602 A; 352 C; 426 G; 482 T; 0 U; 0 Other;
Query Match 13.5%; Score 658.6; DB 4; Length 1862;
Best Local Similarity 99.3%; Pred. No. 3.9e-128;
Matches 672; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 3869 CAGGTTAATCATGATCCAGAAACAGTGTCTTAAGAAGACGCGTGGATGTGTA 3928
DB 1182 CTGGTTAATCATGATCCAGAAACAGTGTCTTAAGAAGACGCGTGGATGTGTA 1241
QY 3929 TAATAGATGTCTATGAGCAATCCAAACGAGATGATCTACTGGGTGGACAGTACACCTG 3988
DB 1242 TAATAGATGTCTATGAGCAATCCAAACGAGATGATCTACTGGGTGGACAGTACACCTG 1301
QY 3989 GGACATGCCAAGCATGCGACAGATGATGGTGTAGTATGATGAATGGAGGGGTGATG 4048
DB 1302 GGACATGCCAAGCATGCGACAGATGATGGTGTAGTATGATGAATGGAGGGGTGATG 1361
QY 4049 GTACTCAATGAGGAGATGATGATCAAGATCAGGCGCTCTCCACAGCAATAGTCCCC 4108
DB 1362 GTACTCAATGAGGAGATGATGATCAAGATCAGGCGCTCTCTCCACAGCAATAGTCCCC 1421
QY 4109 AATAGTATGATTTTGTCTCTCTGATGTGACAAATTTTGTACATATGTTATGGAA 4168
DB 1422 AATAGTATGATTTTGTCTCTCTGATGTGACAAATTTTGTACATATGTTATGGAA 1481
QY 4169 TTTTCTTTTATATATATCT 4228
DB 1482 TTTTCTTTTATATATATCT 1541
QY 4229 AAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4288
DB 1542 AAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1601
QY 4289 TCTCTGTCTCAGCCAGAGTAAACAAAAGTATAGTTTTGACAGAGTGTGTGTCAATTT 4348
DB 1602 TCTCTGTCTCA-CCAGAGTAAACAAAAGTATAGTTTTGACAGAGTGTGTGTCAATTT 1660

QY 4349 TCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 4408
DB 1661 TCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1720
QY 4409 TAGGAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACTATAT 4468
DB 1721 TAGGAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACTATAT 1780
QY 4469 TTATTTATGATAGGATCTGTCAAGAAAGAACTTCCAAAAGAGATTATTAATTAACACGACT 4528
DB 1781 TTATTTATGATAGGATCTGTCAAGAAAGAACTTCCAAAAGAGATTATTAATTAACACGACT 1840
QY 4529 CTGTTGCAATTAAGTTAA 4545
DB 1841 CTGTTGCAATTAAGTTAA 1857
RESULT 14
AAK51625
ID AAK51625 standard; cDNA; 1976 BP.
AC AAK51625;
XX
XX 06-NOV-2001 (first entry)
DE Human polynucleotide SEQ ID NO 170.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX Homo sapiens.
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
XX 03-PEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
DR P-PSDB; AAM78492.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 1; Page 906-909; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and


```
Db 1740 TTTTCTTTTCAATATATATTCCTCTAAAACTCTCAAGCAGACGTCGTCACCTTTTG 1799
Qy 4229 AAAAAAGTATAGGATATATTTACATTTAAATAGCACATGATTTCTTTTGTCTTTCTTCATT 4288
Db 1800 AAAAAAGTATAGGATATATTTACATTTAAATAGCACATGATTTCTTTTGTCTTTCTTCATT 1859
Qy 4289 TCTCTTGTCTCACCCCAAGTAACAAAAAGTATAGTTTTGACAGAGTTCGTTGTTTCATAATT 4348
Db 1860 TCTCTTGTCTCA-CCNAGAGTAACNAAAGTATAGTTTTGACAGAGTTCGTTGTTTCATAATT 1918
Qy 4349 TCAGTTCTAGTTGATTCGAGAAATTTCAAATAAGGAAGAGGGGTCTTTTATFCTTTGTCG 4408
Db 1919 TCAGTTCTAGTTGATTCGAGAAATTTCAAATAAGGAAGAGGGGTCTTTTATCCTTTGTCG 1978
Qy 4409 TAGGAAACCATGACGGAAGGAAACACTGATGTTTAAAGTCCCACTTTTAAACTATAT 4468
Db 1979 TAGGAAACCATGACGGAAGGAAACACTGATGTTTAAAGTCCCACTTTTAAACTATAT 2038
Qy 4469 TTATTTATGTAGGATCTGTCAAAGAAAACTTCCAAAAAGATTTATTAATTTAAACCAGACT 4528
Db 2039 TTATTTATGTAGGATCTGTCAAAGAAAACTTCCAAAAAGATTTATTAATTTAAACCAGACT 2098
Qy 4529 CTGTTGCAATAAGTTAA 4545
Db 2099 CTGTTGCAATAAAAAA 2115
```

Search completed: November 22, 2004, 17:23:20
Job time : 2173.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:21 ; Search time 20190.7 Seconds
(without alignments)
11427.340 Million cell updates/sec

Title: US-10-017-724-5_COPY_4000_8878

Perfect score: 4879

Sequence: 1 tcagaatagtggtgatgag.....gattgcattgagccaagatc 4879

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4879	100.0	8878	6	I47706	I47706 Sequence 3
2	4879	100.0	8878	9	HUMFIBRB	M64983 Homo sapien
3	4879	100.0	8878	11	G28617	G28617 SHGC-35791
4	4783	98.0	156442	9	AC107385	AC107385 Homo sapi
5	4779.4	98.0	11956	9	AF388026	AF388026 Homo sapi
6	2332.4	47.8	330999	2	AC009583	AC009583 Homo sapi
7	2092.4	42.9	330999	2	AC009583	AC009583 Homo sapi
8	2081.2	42.7	4043	6	AX535001	AX535001 Sequence
9	971.8	19.9	4967	6	CQ414791	CQ414791 Sequence
10	971.8	19.9	4967	6	AX285153	AX285153 Sequence
11	696.2	14.3	223534	2	AC140369	AC140369 Mus muscu
12	696.2	14.3	227600	10	AC138394	AC138394 Mus muscu
13	683	14.0	763	9	HUMFBR3	J00132 Human fibri
14	655.4	13.4	1792	9	BC036007	BC036007 Homo sapi
15	653.4	13.4	1883	9	HUMFBRB	J00129 Human fibri
16	625.4	12.8	142856	2	AC143097	AC143097 Macaca mu
17	363.8	7.5	241740	2	AC111892	AC111892 Rattus no
18	363.8	7.5	261110	2	AC105842	AC105842 Rattus no
19	363.8	7.5	263108	2	AC112034	AC112034 Rattus no

C	20	362.2	7.4	571	6	AX410537	AX410537 Sequence
C	21	338.2	6.9	375	6	AX410554	AX410554 Sequence
C	22	337	6.9	384	11	G25542	G25542 human STS E
	23	336.4	6.9	37092	2	AC016278	AC016278 Homo sapi
	24	320.8	6.6	797	9	HUMFBRBB	J00133 human fibri
	25	288.2	5.9	1229	6	CQ723670	CQ723670 Sequence
	26	283.8	5.8	293	11	G05688	G05688 human STS W
	27	283.8	5.8	293	11	G33709	G33709 human STS S
C	28	264.4	5.4	357	6	AX408104	AX408104 Sequence
	29	246.2	5.0	251	9	AF200354S3	AF200356 Pan trogl
	30	246	5.0	261	10	RATFIBBA	M35602 Rat beta-fi
	31	233	4.8	708	10	AF413205	AF413205 Mus muscu
	32	233	4.8	1602	10	BC031715	BC031715 Mus muscu
	33	233	4.8	1612	10	RNU05675	U05675 Rattus norv
	34	230.2	4.7	1329	4	BTPIBR	V00110 Bovine mess
	35	230	4.7	851	9	BC070245	BC070245 Homo sapi
	36	228.6	4.7	380	6	AX409357	AX409357 Sequence
	37	228	4.7	1007	10	RATFIBBB	M27220 Rat fibri
	38	200.4	4.1	1763	5	CHKFIBBC	M58514 Gallus gall
	39	181.4	3.7	1875	5	BC054326	BC054326 Xenopus l
	40	181.4	3.7	2390	5	XLUL9618	U19618 Xenopus lae
	41	180	3.7	230514	2	AC106233	AC106233 Rattus no
	42	162	3.3	162	11	G54120	G54120 1327 Human
C	43	157.8	3.2	181272	5	BX548256	BX548256 Zebrafish
C	44	157.2	3.2	190325	9	AL732327	AL732327 Human DNA
C	45	156.6	3.2	3556	9	AF230338	AF230338 Homo sapi

ALIGNMENTS

RESULT 1
I47706
LOCUS I47706 8878 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 3 from patent US 5639940.
ACCESSION I47706
VERSION I47706.1 GI:2471671
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8878)
AUTHORS Garner, I., Dallymple, M.L., Prunkard, D.E. and Foster, D.C.
TITLE Production of fibrinogen in transgenic animals
JOURNAL Patent: US 5639940-A 3 17-JUN-1997;
FEATURES
source
1..8878
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 100.0%; Score 4879; DB 6; Length 8878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAGAAATAGTCTGATGAGTTAAATAACAATGCGAAGCTGTTCCAGAGCTCTCTT 60
Db 4000 TCAGAAATAGTCTGATGAGTTAAATAACAATGCGAAGCTGTTCCAGAGCTCTCTT 4059
QY 61 CTTCTTTTCAGTACATGATGATTTGCTGAAAGACCTGTGGCAAAAGAGGAGCAAGTAA 120
Db 4060 CTTCTTTTCAGTACATGATGATTTGCTGAAAGACCTGTGGCAAAAGAGGAGCAAGTAA 4119
QY 121 AAGGTAGATATCCTTGTGCTTCCATTCGATTTTCAGCTATATAAATTTGGAACCGTTAGAC 180
Db 4120 AAGGTAGATATCCTTGTGCTTCCATTCGATTTTCAGCTATATAAATTTGGAACCGTTAGAC 4179
QY 181 TGCCACGAGAAATGATGTTGTGAGAGATTAACATTTCTGGGTTAGTGAATGAGATTCA 240
Db 4180 TGCCACGAGAAATGATGTTGTGAGAGATTAACATTTCTGGGTTAGTGAATGAGATTCA 4239
QY 241 TACGGCTTTGGGCACCTTCCCTCCCTGCAACTTGCAGATAGCACTATTTCAGCTTTATTCC 300

Db 4240 TACGCTTTTGGGACCTTCCCTGCAACTTGCAGATAAGCACTATTTCAGCTCTTATTCC 4299
Qy 301 CAGCTGACATCAGCAAGTGTGATTTCTATGAAAAATTTCTACTATGACTCCTTATTTTA 360
Db 4300 CAGCTGACATCAGCAAGTGTGATTTCTATGAAAAATTTCTACTATGACTCCTTATTTTA 4359
Qy 361 AGTATACAAGAACTTTGTGACTCAGAGATAATATTACAGAGTGGAAAAACCCCTAG 420
Db 4360 AGTATACAAGAACTTTGTGACTCAGAGATAATATTACAGAGTGGAAAAACCCCTAG 4419
Qy 421 CATTTATAGTTTAACTTTGAGTTTGAATGAGAGAGTTATCCATATATATTAATTTCAAT 480
Db 4420 CATTTATAGTTTAACTTTGAGTTTGAATGAGAGAGTTATCCATATATATTAATTTCAAT 4479
Qy 481 GTGTTGTGATAATGACACCTTAACCTGTGAATCTTTGAGTCAAGATGTTGAGTGTGTTG 540
Db 4480 GTGTTGTGATAATGACACCTTAACCTGTGAATCTTTGAGTCAAGATGTTGAGTGTGTTG 4539
Qy 541 ACTTGGTGGTCAGGAAACAGCTAGTGCCTGAGCCTGGCAAGGCACTCTCAGTGAGTAGCA 600
Db 4540 ACTTGGTGGTCAGGAAACAGCTAGTGCCTGAGCCTGGCAAGGCACTCTCAGTGAGTAGCA 4599
Qy 601 TACCACAGTTGGAAATTTTCAAGAAATCAAGGAATCATGACATCTTATAAATTTCA 660
Db 4600 TACCACAGTTGGAAATTTTCAAGAAATCAAGGAATCATGACATCTTATAAATTTCA 4659
Qy 661 AGGTTCTGCTATACCTTATGTGAATGGATAAATAATCAAGCATATCAACTCTGTAAGAT 720
Db 4660 AGGTTCTGCTATACCTTATGTGAATGGATAAATAATCAAGCATATCAACTCTGTAAGAT 4719
Qy 721 TGAATCTCTCAGATGGAAGACCCCAATATCTGCTTTCTCTCTTTTCCCTCACCAAGAAA 780
Db 4720 TGAATCTCTCAGATGGAAGACCCCAATATCTGCTTTCTCTCTTTTCCCTCACCAAGAAA 4779
Qy 781 TAAACAACTATTTCATTTATCTGGAACAATCTTTAGCGTATACCTATGTTGTAATTA 840
Db 4780 TAAACAACTATTTCATTTATCTGGAACAATCTTTAGCGTATACCTATGTTGTAATTA 4839
Qy 841 CTAGTATGCTGTAGGATTTATGTTAATTTGTAATGATGATGATGATGATGATGATGAT 900
Db 4840 CTAGTATGCTGTAGGATTTATGTTAATTTGTAATGATGATGATGATGATGATGATGAT 4899
Qy 901 TAAATATGACTATATATATCTGCTTGGTGATAGCTCAGTGTGTTAATAGTTTATCTC 960
Db 4900 TAAATATGACTATATATATCTGCTTGGTGATAGCTCAGTGTGTTAATAGTTTATCTC 4959
Qy 961 AGAAATCAAAATGCTATAGTTAATAATATAGTTTATGAGGCAAAATGCTACTATT 1020
Db 4960 AGAAATCAAAATGCTATAGTTAATAATATAGTTTATGAGGCAAAATGCTACTATT 5019
Qy 1021 TCTACATAATTTCAATTTTCCAGATAATGAAATGATGATGATGATGATGATGATGATGAT 1080
Db 5020 TCTACATAATTTCAATTTTCCAGATAATGAAATGATGATGATGATGATGATGATGATGAT 5079
Qy 1081 GGAAAGCACCAATATATATAGATGAGCTGTGTAATAGCAATATCCCAACTAACCTTCG 1140
Db 5080 GGAAAGCACCAATATATATAGATGAGCTGTGTAATAGCAATATCCCAACTAACCTTCG 5139
Qy 1141 TGTGCTGCTTCAATCTGGAACCTGAGAGCAAAATACAAAGTTAGATCTGATGT 1200
Db 5140 TGTGCTGCTTCAATCTGGAACCTGAGAGCAAAATACAAAGTTAGATCTGATGT 5199
Qy 1201 CTCAGCTCAAAATGGAATATGTCGACACCCCATGCTGTCAGTTCAATATTTCTGTGTGT 1260
Db 5200 CTCAGCTCAAAATGGAATATGTCGACACCCCATGCTGTCAGTTCAATATTTCTGTGTGT 5259
Qy 1261 GTCTGGCAAGGTAACTGATTTCAATAATATTTTATGAGAGTTCCAGAGACTCACAC 1320
Db 5260 GTCTGGCAAGGTAACTGATTTCAATAATATTTTATGAGAGTTCCAGAGACTCACAC 5319
Qy 1321 ACCAAAAATAAGAGAACCAACCAACCAAAAAATGCTAAGTGGATTTTCCCAACAGATC 1380
Db 5320 ACCAAAAATAAGAGAACCAACCAACCAAAAAATGCTAAGTGGATTTTCCCAACAGATC 5379

Qy 1381 ATAATGACATTTACAGTACATCATATAAATAATCCTTAGCCAGTTGTGTTTGGACTGGCCT 1440
Db 5380 ATAATGACATTTACAGTACATCATATAAATAATCCTTAGCCAGTTGTGTTTGGACTGGCCT 5439
Qy 1441 GGTGCATTTGCTGTTTGTATGACAGGATGGGGCACAGGTAGTGTCCCAAGGGTGGCTGAT 1500
Db 5440 GGTGCATTTGCTGTTTGTATGACAGGATGGGGCACAGGTAGTGTCCCAAGGGTGGCTGAT 5499
Qy 1501 GTGTGCATCTCGTACTGGCTTTGAACAGATGGCAGAACCAAGATAGATGTAGAGTTTC 1560
Db 5500 GTGTGCATCTCGTACTGGCTTTGAACAGATGGCAGAACCAAGATAGATGTAGAGTTTC 5559
Qy 1561 TCCATTTTGTGTGTTCTGGGAGCTCATGGAATATTCAGGAACAACAAAGGTGGAGAGAGC 1620
Db 5560 TCCATTTTGTGTGTTCTGGGAGCTCATGGAATATTCAGGAACAACAAAGGTGGAGAGAGC 5619
Qy 1621 TTTGTTCACTCTCTTGTAGCAGATAAAGCTCTCAAACTGGGTTGGACTTAAAGTAAA 1680
Db 5620 TTTGTTCACTCTCTTGTAGCAGATAAAGCTCTCAAACTGGGTTGGACTTAAAGTAAA 5679
Qy 1681 ATGAAATCTAATATTTGTTATATTTTCAAGGTCTATAAACAACACTCCTTAGTA 1740
Db 5680 ATGAAATCTAATATTTGTTATATTTTCAAGGTCTATAAACAACACTCCTTAGTA 5739
Qy 1741 ACTTATGTAAATGTTTATTTTAAAGAAATGCTGACTTAAATACAAAGTAAATATGTCTATAAAC 1800
Db 5740 ACTTATGTAAATGTTTATTTTAAAGAAATGCTGACTTAAATACAAAGTAAATATGTCTATAAAC 5799
Qy 1801 CCCTGAAACATTAATGTTGTTTACATTTGCGAATGTGAGGAAAATTTATCAGGAAAGAGGT 1860
Db 5800 CCCTGAAACATTAATGTTGTTTACATTTGCGAATGTGAGGAAAATTTATCAGGAAAGAGGT 5859
Qy 1861 GAAACATCTGAAATGTAATCTCATCACTCAAGTCTGTCMAACCGTATAGAGTATAC 1920
Db 5860 GAAACATCTGAAATGTAATCTCATCACTCAAGTCTGTCMAACCGTATAGAGTATAC 5919
Qy 1921 TGTGACATGAATACAGAAATGAGAGTAAGCTTTTCGACAGTTGTTGACTGTTGATCTGT 1980
Db 5920 TGTGACATGAATACAGAAATGAGAGTAAGCTTTTCGACAGTTGTTGACTGTTGATCTGT 5979
Qy 1981 AATTAATTTGGATACCGTAAATGCGAGAAACAGAGCCAGGTGTGGTGGCTCATACCTGT 2040
Db 5980 AATTAATTTGGATACCGTAAATGCGAGAAACAGAGCCAGGTGTGGTGGCTCATACCTGT 6039
Qy 2041 AATTCACGACCTTGGGAGGCCAAAGTGGCTGATGCTGAGCCTAGAGTTTCAAACT 2100
Db 6040 AATTCACGACCTTGGGAGGCCAAAGTGGCTGATGCTGAGCCTAGAGTTTCAAACT 6099
Qy 2101 AGCCTGGGCAACATAATGAGACCCCTAACTCTACAAAAAATAAATAACCAAAAAAAA 2160
Db 6100 AGCCTGGGCAACATAATGAGACCCCTAACTCTACAAAAAATAAATAACCAAAAAAAA 6159
Qy 2161 AAAAAATCAGCTGTGTTGGTAGTATGCTGTAGTCCAGCTATCCAGGAGGTGAGA 2220
Db 6160 AAAAAATCAGCTGTGTTGGTAGTATGCTGTAGTCCAGCTATCCAGGAGGTGAGA 6219
Qy 2221 TGGGAGATCAGCTGAGCCCAACCTGAGTCTTGTATCTGCTAGTCTGAGCTAGCCTGG 2280
Db 6220 TGGGAGATCAGCTGAGCCCAACCTGAGTCTTGTATCTGCTAGTCTGAGCTAGCCTGG 6279
Qy 2281 GCAACAGAGGATAGTGTGATCCTGCTCAAAAAAATAAATAAATAAAGCCAGGAAAC 2340
Db 6280 GCAACAGAGGATAGTGTGATCCTGCTCAAAAAAATAAATAAATAAAGCCAGGAAAC 6339
Qy 2341 AAGACTTAGTCTTAACATCTAAGTGTGAGGAGTAAATTTGATGCTGGAATTCAC 2400
Db 6340 AAGACTTAGTCTTAACATCTAAGTGTGAGGAGTAAATTTGATGCTGGAATTCAC 6399
Qy 2401 CTGATTTTAAAGTTATAAATAATCTATAATTTCAAAATTTGGGGTAAAGATAAGACATT 2460
Db 6400 CTGATTTTAAAGTTATAAATAATCTATAATTTCAAAATTTGGGGTAAAGATAAGACATT 6459


```

Db      8620 CTCAATAGCTAATAAACCGGCTTGAATATTTGAAGATTTAAATCTGACTCTAGGACG 8679
Qy      4681 GGCAGGTGGCTCAGACTATAATCCCAACACTTTGGGAGGCTGAGGCGGGTCAAA 4740
Db      8680 GGCAGGTGGCTCAGACTATAATCCCAACACTTTGGGAGGCTGAGGCGGGTCAAA 8739
Qy      4741 GGTCAAGAGTTCAAGACGAGCTGACCAATATGGTGAACCCCATCTCTACTAAAAATAC 4800
Db      8740 GGTCAAGAGTTCAAGACGAGCTGACCAATATGGTGAACCCCATCTCTACTAAAAATAC 8799
Qy      4801 AAAAATAGCCAGGCGTGGTGGCAGTCCCTGTAGTCCAGCTAGCCTGTGAGGTGAG 4860
Db      8800 AAAAATAGCCAGGCGTGGTGGCAGTCCCTGTAGTCCAGCTAGCCTGTGAGGTGAG 8859
Qy      4861 ATTGCATTGAGCAAGATC 4879
Db      8860 ATTGCATTGAGCAAGATC 8878

RESULT 2
LOCUS   HUMFIBRB                      8878 bp    DNA          linear    PRI 18-MAY-2000
DEFINITION Homo sapiens fibrinogen beta chain (FGB), complete cds.
ACCESSION M64983
VERSION   M64983.1 GI:182597
KEYWORDS .
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 8878)
AUTHORS   Chung,D.W., Harris,J.E. and Davie,E.W.
TITLE     Nucleotide sequences of the three genes coding for human fibrinogen
JOURNAL   (in) Liu,C.Y. and Chien,S. (Eds.);
FIBRINOGEN, THROMBOSIS, COAGULATION AND FIBRINOLYSIS: 39-48;
Plenum Press, New York (1991)
FEATURES             Location/Qualifiers
     source           1..8878
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene             <470..8878
                     /gene="FGB"
     mRNA             join(<470..583,3258..3449,3939..4122,5043..5270,
                     5831..5944,6633..6758,6967..7252,7871..8537)
                     /gene="FGB"
     mRNA             /product="fibrinogen beta chain"
                     join(<470..583,3258..3449,3939..4122,5043..5270,
                     5831..5944,6633..6758,6967..7252,7871..8534)
                     /gene="FGB"
     mRNA             /product="fibrinogen beta chain"
                     join(470..583,3258..3449,3939..4122,5043..5270,5831..5944,
                     6633..6758,6967..7252,7871..8273)
                     /gene="FGB"
     mRNA             /product="fibrinogen beta chain"
                     join(<470..583,3258..3449,3939..4122,5043..5270,
                     5831..5944,6633..6758,6967..7252,7871..8268)
                     /gene="FGB"
     mRNA             /product="fibrinogen beta chain"
                     join(<470..583,3258..3449,3939..4122,5043..5270,
                     5831..5944,6633..6758,6967..7252,7871..8202)
                     /gene="FGB"
     CDS             /product="fibrinogen beta chain"
                     join(470..583,3258..3449,3939..4122,5043..5270,5831..5944,
                     6633..6758,6967..7252,7871..8102)
                     /gene="FGB"
                     /note="start codon not confirmed"
                     /codon_start=1
                     /product="fibrinogen beta chain"

```

```

/protein_id="AAA18024.2"
/db_xref="GI:7924018"
/translation="MKRMVSWSFHKLTKMHLKLLLCVFLVKQGVNDNEEGFSPAR
GHRPLDKRREAPSLRPAPPISGGYRARPAAATOKKVERKAPDAGGCLHADPDL
GVLCTQCQLQELLOQERPIRNSVDLNNNVEAVSQTSSSFQYMYLLKDLWQKROK
QKPTNNVNEYSESLBKHLQYIDETVNNNIPNLRLVLRSLRSLKQLGSDVSA
QMETVPIQNRDQGSVDFGRKMDPYKQGFNGFVATNTDKNYCGLPGYVWLGNDKIQLTR
MGPTLLEIMEDWKGDKVKAHYGGFTVQNEANKYQISVNYRGTTAGNALMDGASLGMG
ENRPMTHNGMFFSYDNDNGLTSDPKCKSKEDGGGWWYNPCHAANPNRYVMG
QYTWDMAKHGIDDGVMWNWKGWSYMRKMSMKIRPFPQQ"
8103..8537
/gene="FGB"
8538..8878
/gene="FGB"
/note="3' flanking sequence"

3'UTR
misc_feature
ORIGIN
Query Match      100.0%; Score 4879; DB 9; Length 8878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCAGACCTCTCTT 60
Db      4000 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCAGACCTCTCTT 4059
Qy      61 CTTCCTTTTCAGTACATGATTTTGTCTGAAAGACCTGTGGCAAAAGAGGAGCAAGTAA 120
Db      4060 CTTCCTTTTCAGTACATGATTTTGTCTGAAAGACCTGTGGCAAAAGAGGAGCAAGTAA 4119
Qy      121 AAGTAGATATCTTGTGCTTTTCCATTTCGATTTTCAGCTATAAAATTGGAACCGTTAGAC 180
Db      4120 AAGTAGATATCTTGTGCTTTTCCATTTCGATTTTCAGCTATAAAATTGGAACCGTTAGAC 4179
Qy      181 TGCCACGAGATGATGCTGCTGTGAGCAATTAACAATTTCTGGTGTAGTGAATAGCATCA 240
Db      4180 TGCCACGAGATGATGCTGCTGTGAGCAATTAACAATTTCTGGTGTAGTGAATAGCATCA 4239
Qy      241 TACGCTTTTGGGCACCTTCCCTGCAACTTGCCAGATAAGCACTATTTCAGCTCTTATTCC 300
Db      4240 TACGCTTTTGGGCACCTTCCCTGCAACTTGCCAGATAAGCACTATTTCAGCTCTTATTCC 4299
Qy      301 CAGTCTGATCAGCAAGTGTGATTTTCTATGAAAAATTTCTATGACTCTCTTATTTTA 360
Db      4300 CAGTCTGATCAGCAAGTGTGATTTTCTATGAAAAATTTCTATGACTCTCTTATTTTA 4359
Qy      361 AGTATACAGAAACTTGTGACTCAGAGATATATTTACAGAGTGGAAAAAACCCCTAG 420
Db      4360 AGTATACAGAAACTTGTGACTCAGAGATATATTTACAGAGTGGAAAAAACCCCTAG 4419
Qy      421 CATTTATAGTTTAAACATTTGAGGTTTGAATGAGAGAGTTATCCATATATATTAATT 480
Db      4420 CATTTATAGTTTAAACATTTGAGGTTTGAATGAGAGAGTTATCCATATATATTAATT 4479
Qy      481 GTGTTGTGGATAATGACACCTAACTGTGAAATCTTGGAGTTCAGATGTTGAGTGTCTT 540
Db      4480 GTGTTGTGGATAATGACACCTAACTGTGAAATCTTGGAGTTCAGATGTTGAGTGTCTT 4539
Qy      541 ACTTGTGCTCAGGAAACAGCTAGTGGCTGAGCCTGGCAGCAGGCACTCTCAGTGAGTACA 600
Db      4540 ACTTGTGCTCAGGAAACAGCTAGTGGCTGAGCCTGGCAGCAGGCACTCTCAGTGAGTACA 4599
Qy      601 TACCACAGTTGGAAATTTTCAAAGAAATCAAAGAAATCAATCATCTTATAAATTTCA 660
Db      4600 TACCACAGTTGGAAATTTTCAAAGAAATCAAAGAAATCAATCATCTTATAAATTTCA 4659
Qy      661 AGGTTCTGCTATCTTATGTAATGGAATAATAAATCAAGCATATCACTCTGTGAAGAT 720
Db      4660 AGGTTCTGCTATCTTATGTAATGGAATAATAAATCAAGCATATCACTCTGTGAAGAT 4719
Qy      721 TGAATCTCTCAGATGGAAGAGCCCAATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db      4720 TGAATCTCTCAGATGGAAGAGCCCAATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 4779

```

Qy	781	TAAACAACCTATTTCATTTTATTACTGGACACAATCTTTAGCGTATACCTATGGTAAATTA	840
Db	4780	TAAACAACCTATTTCATTTTATTACTGGACACAATCTTTAGCGTATACCTATGGTAAATTA	4839
Qy	841	CTAGTAGTGGTGGTTAGGATTTATGTTAAATTTGTATATGTCATGCGCCAAATCATTTCCAC	900
Db	4840	CTAGTAGTGGTGGTTAGGATTTATGTTAAATTTGTATATGTCATGCGCCAAATCATTTCCAC	4899
Qy	901	TAAATATGACTATATATCATAACTGCTTGGTGATPAGCTCAGTGTGTTTAAATAGTTTATTTCTC	960
Db	4900	TAAATATGACTATATATCATAACTGCTTGGTGATPAGCTCAGTGTGTTTAAATAGTTTATTTCTC	4959
Qy	961	AGAAAATCAAAATTTGTTAGTTAAATACATTTAGTTTATGAGGCAAAAATGCTAACTATT	1020
Db	4960	AGAAAATCAAAATTTGTTAGTTAAATACATTTAGTTTATGAGGCAAAAATGCTAACTATT	5019
Qy	1021	TCTACATAAATTTCAATTTTCCAGATAATGAAAATGTAGTCAATGAGTACTCCTCAGAACT	1080
Db	5020	TCTACATAAATTTCAATTTTCCAGATAATGAAAATGTAGTCAATGAGTACTCCTCAGAACT	5079
Qy	1081	GGAAAAGCACCAATTTATATATAGATGAGACTGTGAATAGCAATATCCAACTAACCTTCG	1140
Db	5080	GGAAAAGCACCAATTTATATATAGATGAGACTGTGAATAGCAATATCCAACTAACCTTCG	5139
Qy	1141	TGTGCTTCGTTCAATCTCGGAAACCTGGAAGCAAAATACAAAAGTTAGAACTGATGT	1200
Db	5140	TGTGCTTCGTTCAATCTCGGAAACCTGGAAGCAAAATACAAAAGTTAGAACTGATGT	5199
Qy	1201	CTCAGCTCAAATGGAAATATTGTGCGACCCCATGCACTGTCAAGTTGCAATATTTCTGTGGT	1260
Db	5200	CTCAGCTCAAATGGAAATATTGTGCGACCCCATGCACTGTCAAGTTGCAATATTTCTGTGGT	5259
Qy	1261	GTCTGGCAAAAGGTAACGTATCAATAACATATTTTATAGAGTTCCAGAGAACTCACAC	1320
Db	5260	GTCTGGCAAAAGGTAACGTATCAATAACATATTTTATAGAGTTCCAGAGAACTCACAC	5319
Qy	1321	ACCAAAATTAAGAGAACAAACAAACAAACAAATGCTAAGTGATTTTCCCAACAGATC	1380
Db	5320	ACCAAAATTAAGAGAACAAACAAACAAACAAATGCTAAGTGATTTTCCCAACAGATC	5379
Qy	1381	ATAATGACATTACAGTACATCAAAAAATATCCTTAGCCAGTTGTGTTTGGACTGGCCT	1440
Db	5380	ATAATGACATTACAGTACATCAAAAAATATCCTTAGCCAGTTGTGTTTGGACTGGCCT	5439
Qy	1441	GGTGCAATTTGCTGGTTTGTATGAGCAGATGGGCGACAGGTAGTCCGAGGGTGGCTGAT	1500
Db	5440	GGTGCAATTTGCTGGTTTGTATGAGCAGATGGGCGACAGGTAGTCCGAGGGTGGCTGAT	5499
Qy	1501	GTGTGCATCTCGGTACTGGCTTGAAACAGATGGCAGAACCAACAGATAGATGTAGAAGTTTC	1560
Db	5500	GTGTGCATCTCGGTACTGGCTTGAAACAGATGGCAGAACCAACAGATAGATGTAGAAGTTTC	5559
Qy	1561	TCCATTTTGTGTTCTGGGAGCTCATGGATATTTCCAGGACACAAAAAGGTGGAGAAGGC	1620
Db	5560	TCCATTTTGTGTTCTGGGAGCTCATGGATATTTCCAGGACACAAAAAGGTGGAGAAGGC	5619
Qy	1621	TTTGTTCATCCTCTTAGCAGATAAACGTCCTCAAACTGGGGTGGACTTACTAAAGTAAA	1680
Db	5620	TTTGTTCATCCTCTTAGCAGATAAACGTCCTCAAACTGGGGTGGACTTACTAAAGTAAA	5679
Qy	1681	ATGAAAATCTAAATTTTGTTATATTTTCAAAGGTCATAATAACACACTCCTTAGTA	1740
Db	5680	ATGAAAATCTAAATTTTGTTATATTTTCAAAGGTCATAATAACACACTCCTTAGTA	5739
Qy	1741	ACTTATGTAATGTTTATTTTAAAGAAATGGTGACTTAAATACAAAAGTAAATATGTCTAAAC	1800
Db	5740	ACTTATGTAATGTTTATTTTAAAGAAATGGTGACTTAAATACAAAAGTAAATATGTCTAAAC	5799
Qy	1801	CCCTGGAACATAAGTTGTCCTTACATTTGCGAGATGTGAGGAAATTATCAGAAAGGAGGT	1860
Db	5800	CCCTGGAACATAAGTTGTCCTTACATTTGCGAGATGTGAGGAAATTATCAGAAAGGAGGT	5859

Qy	1861	GAACA	CTCTGAA	ATGTATCTCA	TTCAA	CTGTGACAG	ATTTCTGT	CCAA	CCGTAT	PAGAGTATAC	1921
Db	5860	GAACA	CTCTGAA	ATGTATCTCA	TTCAA	CTGTGTGACAG	ATTTCTGT	CCAA	CCGTAT	PAGAGTATAC	5919
Qy	1921	TGTGAC	ATGAATACAG	AAAAATGGAGG	TAAGCTTT	CGACAGCTTG	TTGACCTG	TGTTGACCTG	TGTTGATCTGT	1980	
Db	5920	TGTGAC	ATGAATACAG	AAAAATGGAGG	TAAGCTTT	CGACAGCTTG	TTGACCTG	TGTTGACCTG	TGTTGATCTGT	5979	
Qy	1981	AAATTA	TTTGGATACCG	TAAAAATCCAGGA	AAACAAGCCAGG	CTGTGTG	TGGCTCATAC	CTCTGT	2040		
Db	5980	AAATTA	TTTGGATACCG	TAAAAATCCAGGA	AAACAAGCCAGG	CTGTGTG	TGGCTCATAC	CTCTGT	6039		
Qy	2041	AAATTC	CAGACCTTTGGG	AGGCCAAAGTGGG	CTGATAGCTTTGAGC	CTTAGGAGT	TTGAAACT	2100			
Db	6040	AAATTC	CAGACCTTTGGG	AGGCCAAAGTGGG	CTGATAGCTTTGAGC	CTTAGGAGT	TTGAAACT	6099			
Qy	2101	AGCCTGGG	CAACATAATGAG	ACCCTACTCTACAAAA	AAAAAATAATCCAAAA	AAAAA	2160				
Db	6100	AGCCTGGG	CAACATAATGAG	ACCCTACTCTACAAAA	AAAAAATAATCCAAAA	AAAAA	6159				
Qy	2161	AAAAAA	ATCAGCTGTGT	TGTTAGTATGTG	CTGTGTAGTCCAG	CTATCCAGGAGC	CTGAGA	2220			
Db	6160	AAAAAA	ATCAGCTGTGT	TGTTAGTATGTG	CTGTGTAGTCCAG	CTATCCAGGAGC	CTGAGA	6219			
Qy	2221	TGGGAG	ATCAGCTGAG	CCCAACAACCTGG	AGTCTTGATCATG	TACTGTAAC	CTGTAGCCTG	2280			
Db	6220	TGGGAG	ATCAGCTGAG	CCCAACAACCTGG	AGTCTTGATCATG	TACTGTAAC	CTGTAGCCTG	6279			
Qy	2281	GCACAG	AGATAGTGAG	ATCCTGTCTCAAAAA	AAAAAATAATTA	AAAAAGCCAGG	AAAC	2340			
Db	6280	GCACAG	AGATAGTGAG	ATCCTGTCTCAAAAA	AAAAAATAATTA	AAAAAGCCAGG	AAAC	6339			
Qy	2341	AAGACT	TTAGCTTAA	CACTCTAACATAG	CTGACAAAGGAG	TAAATGATG	TGGAATCAAC	2400			
Db	6340	AAGACT	TTAGCTTAA	CACTCTAACATAG	CTGACAAAGGAG	TAAATGATG	TGGAATCAAC	6399			
Qy	2401	CTGATAT	TTAAAAAGTTA	AAAAATCTATAAT	TTCACAAATTTGGG	GTAAAGATAAG	ACACTT	2460			
Db	6400	CTGATAT	TTAAAAAGTTA	AAAAATCTATAAT	TTCACAAATTTGGG	GTAAAGATAAG	ACACTT	6459			
Qy	2461	GCAGTT	TTCCAAAGTTTTAC	AGTTTTACCTCTCATAT	TATTTTCCTATTTG	TGTCCTATTT	2520				
Db	6460	GCAGTT	TTCCAAAGTTTTAC	AGTTTTACCTCTCATAT	TATTTTCCTATTTG	TGTCCTATTT	6519				
Qy	2521	TAGAGC	ACCAATATATAC	TAAATGGAATGGAC	AGGGGATTCAGATAT	TATTTTCAAAGT	2580				
Db	6520	TAGAGC	ACCAATATATAC	TAAATGGAATGGAC	AGGGGATTCAGATAT	TATTTTCAAAGT	6579				
Qy	2581	GACATTA	TTTGTGTTGTTAA	TATATATGCTTTTTT	GTCTGTCTGTCA	CCAAAGGATGGAC	2640				
Db	6580	GACATTA	TTTGTGTTGTTAA	TATATATGCTTTTTT	GTCTGTCTGTCA	CCAAAGGATGGAC	6639				
Qy	2641	AGTGAT	TTCAAA	CCGTCAAGCGGTAG	TGTGACTTTGGCAGAA	ATGGGATCCATATA	2700				
Db	6640	AGTGAT	TTCAAA	CCGTCAAGCGGTAG	TGTGACTTTGGCAGAA	ATGGGATCCATATA	6699				
Qy	2701	ACAGGG	ATTTGGAAATGTT	CCAACCAACAGAT	GGGAAGAAATTA	CTGTGGCCTAC	CAGG	2760			
Db	6700	ACAGGG	ATTTGGAAATGTT	CCAACCAACAGAT	GGGAAGAAATTA	CTGTGGCCTAC	CAGG	6759			
Qy	2761	TAA	CGAACAGGCAT	GCAAAAATAAAAT	CATTCTATTTGAAAT	TGGGATTTTTTTTT	TAATATAA	2820			
Db	6760	TAA	CGAACAGGCAT	GCAAAAATAAAAT	CATTCTATTTGAAAT	TGGGATTTTTTTTT	TAATATAA	6819			
Qy	2821	AAACAT	TCTATTTGGA	AGCCTGTTTTAG	CGAGTTAAGAGG	AGTTTCTCTGAC	AAAAATGT	2880			
Db	6820	AAACAT	TCTATTTGGA	AGCCTGTTTTAG	CGAGTTAAGAGG	AGTTTCTCTGAC	AAAAATGT	6879			
Qy	2881	GGAAGC	TAAAGATAAGG	GAAGAAAGG	CGTTTTTGT	TTTCCCAAAATTTAT	TTTTTGGTG	2940			
Db	6880	GGAAGC	TAAAGATAAGG	GAAGAAAGG	CGTTTTTGT	TTTCCCAAAATTTAT	TTTTTGGTG	6939			
Qy	2941	AGAGAT	TTTTATTTG	TTTTTTCTTTT	TAGTGAATAT	TGGGCTTGGAAT	TGATAAAATTAGCC	3000			

JOURNAL
COMMENT

Unpublished (2000)

Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: olivier@shgc.stanford.edu
Primer A: TCCCAATACAGTATTTTTCG
Primer B: ACTTTGTACTCTTGTGGTGAGC
STS size: 215
PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:

Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:

MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from M64983
-- Washington University/Merck EST sequence.

FEATURES

source
Location/Qualifiers
1..8878
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map_id="4"
/clone_lib="Human"
8103..8317
8103..8124
complement(8294..8317)
STS
primer_bind
primer_bind
ORIGIN
Query Match 100.0%; Score 4879; DB 11; Length 8878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCGAGACCTCTCTT 60
DB 4000 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCGAGACCTCTCTT 4059
QY 61 CTTCTCTTCAGTACATGATTTTGTCTGAAGACCTGTGGCAAGAGGAGGAGCAAGTAA 120
DB 4060 CTTCTCTTCAGTACATGATTTTGTCTGAAGACCTGTGGCAAGAGGAGGAGCAAGTAA 4119
QY 121 AAGGTAGATATCTTGCTGTTCCATTCGATTTTCAGCTTATAAAATTTGGAACCGTTAGAC 180
DB 4120 AAGGTAGATATCTTGCTGTTCCATTCGATTTTCAGCTTATAAAATTTGGAACCGTTAGAC 4179
QY 181 TCCGAGGAATGCATGGTGTGAGAGAGATTAACTTCTGGGTAGTGAATGCAATTC 240
DB 4180 TCCGAGGAATGCATGGTGTGAGAGAGATTAACTTCTGGGTAGTGAATGCAATTC 4239
QY 241 TAGCCTTTTGGGACCTTCCCTGCAACTTTCAGATAAGCACTATTTCAGCTCTTATCC 300
DB 4240 TAGCCTTTTGGGACCTTCCCTGCAACTTTCAGATAAGCACTATTTCAGCTCTTATCC 4299
QY 301 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAATTTCTACTATGACTCTCTTATTTA 360
DB 5380 ATAATGACATACATACATATAAAATATCCTTAGCCAGTTGTTGTTTGGACTGGCT 5439

DB 4300 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAATTTCTACTATGACTCTCTTATTTA 4359
QY 361 AGTATACAAGAACTTTGTGACTCAGAAGATAATATTACAGAGTGGAAGAAACCCCTAG 420
DB 4360 AGTATACAAGAACTTTGTGACTCAGAAGATAATATTACAGAGTGGAAGAAACCCCTAG 4419
QY 421 CATTATATAGTTTAAACATTTGAGGTTTGAATCAGAGAGTTATCCATAATATATTTCAAT 480
DB 4420 CATTATATAGTTTAAACATTTGAGGTTTGAATCAGAGAGTTATCCATAATATATTTCAAT 4479
QY 481 GTGTTGTGATAATGACACCTAACCTGTGAATCTTGTAGGTGAGAATGTTGAGTGTGTTG 540
DB 4480 GTGTTGTGATAATGACACCTAACCTGTGAATCTTGTAGGTGAGAATGTTGAGTGTGTTG 4539
QY 541 ACTTGTGTGTCAGGAAACAGCTAGTGTGCTGAGCTGGCAAGGCATCTCAGTGAGTAGCA 600
DB 4540 ACTTGTGTGTCAGGAAACAGCTAGTGTGCTGAGCTGGCAAGGCATCTCAGTGAGTAGCA 4599
QY 601 TACCACACAGTTGGAATTTTCAAGAAATCAAGGAATCAAGCAATCTTATAAAATTTCA 660
DB 4600 TACCACACAGTTGGAATTTTCAAGAAATCAAGGAATCAAGCAATCTTATAAAATTTCA 4659
QY 661 AGGTTCTGCTATACCTTATGTGAATGGAATAAATCAAGCATATCCACTCTGTAAGAT 720
DB 4660 AGGTTCTGCTATACCTTATGTGAATGGAATAAATCAAGCATATCCACTCTGTAAGAT 4719
QY 721 TGAACCTTCTCAGATGGAAGACCCCAATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 780
DB 4720 TGAACCTTCTCAGATGGAAGACCCCAATCTGCTTCTCTCTCTCTCTCTCTCTCTCT 4779
QY 781 TAAACAACCTTATTTCAATTTTACTGGACACATCTTTAGCGTATACCTTATGTAATTTA 840
DB 4780 TAAACAACCTTATTTCAATTTTACTGGACACATCTTTAGCGTATACCTTATGTAATTTA 4839
QY 841 CTAGTATGCTGTGTTAGGATTTATGTTAAATTTGTATATGTGTCGCGCAAAATCATTTCCAC 900
DB 4840 CTAGTATGCTGTGTTAGGATTTATGTTAAATTTGTATATGTGTCGCGCAAAATCATTTCCAC 4899
QY 901 TAAATATGACTATATATCATTAATCTGCTTGTGTAAGTCTCAGTGTGTTTAAATGTTATCTC 960
DB 4900 TAAATATGACTATATATCATTAATCTGCTTGTGTAAGTCTCAGTGTGTTTAAATGTTATCTC 4959
QY 961 AGAAAATCAAAATGTATAGTTTAAATACATAGTTTATGAGCAAAATGCTAACTATT 1020
DB 4960 AGAAAATCAAAATGTATAGTTTAAATACATAGTTTATGAGCAAAATGCTAACTATT 5019
QY 1021 TCTACATAATTTCAATTTTCCAGATAATGAAAATGTAGTCAATGAGTACTCTCTCAGAACT 1080
DB 5020 TCTACATAATTTCAATTTTCCAGATAATGAAAATGTAGTCAATGAGTACTCTCTCAGAACT 5079
QY 1081 GGAAAAGCAACCAATTTATATATAGATGAGACTGTGTAATAGCAATATCCCAACTAACCCTCG 1140
DB 5080 GGAAAAGCAACCAATTTATATATAGATGAGACTGTGTAATAGCAATATCCCAACTAACCCTCG 5139
QY 1141 TGTGCTTCTGTTCAATCTCGAAAACTGAGAAGCAAAATACAAAAGTTAGAAATCTGATCT 1200
DB 5140 TGTGCTTCTGTTCAATCTCGAAAACTGAGAAGCAAAATACAAAAGTTAGAAATCTGATCT 5199
QY 1201 CTCAGCTCAAAATGGAATATTTGTCGACCCCACTGCACTGTGCAATATTTCTCTGTGT 1260
DB 5200 CTCAGCTCAAAATGGAATATTTGTCGACCCCACTGCACTGTGCAATATTTCTCTGTGT 5259
QY 1261 GTCTGGCAAGGTAACTGATTCATATAAATATTTTGTAGAGGTTCCAGAGAACTCCACAC 1320
DB 5260 GTCTGGCAAGGTAACTGATTCATATAAATATTTTGTAGAGGTTCCAGAGAACTCCACAC 5319
QY 1321 ACCAAAATTAAGAGAACCAACCAACCAAAATGCTAAAGTGAATTTTCCCAACAGATC 1380
DB 5320 ACCAAAATTAAGAGAACCAACCAACCAAAATGCTAAAGTGAATTTTCCCAACAGATC 5379
QY 1381 ATAATGACATACAGTACATATAAAATATCCTTAGCCAGTTGTTGTTTGGACTGGCT 1440
DB 5380 ATAATGACATACAGTACATATAAAATATCCTTAGCCAGTTGTTGTTTGGACTGGCT 5439

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP13-58007; the clone sequenced to the right is RP11-216A23, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-158C21; actual end is at base position 32435 of RP11-216A23.

Polymorphisms have been identified between AC107385 and AC092616.

FEATURES

source

Location/Qualifiers
1..156442
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-158C21"
/clone_lib="RPCI-11"

misc_feature
1..646
/note="similar to Homo sapiens EST BF205005 (NID:G11098591)"

misc_feature
1..615
/note="similar to Homo sapiens EST BF307717 (NID:G11254872)"

misc_feature
1..535
/note="similar to Homo sapiens EST BG394804 (NID:G13288252)"

misc_feature
1..480
/note="similar to Homo sapiens EST BI868734 (NID:G16042407)"

misc_feature
1..417
/note="similar to Homo sapiens EST BE312794 (NID:G9131943)"

misc_feature
1..337
/note="similar to Homo sapiens EST BF237442 (NID:G11151360)"

misc_feature
1..89
/note="similar to Mus musculus EST BG093388 (NID:G12575951) ut70h03.y1"

misc_feature
1..75
/note="similar to Homo sapiens EST AI637897 (NID:G4690131) tt11f10.x1"

misc_feature
32..765
/note="similar to Homo sapiens EST BG755962 (NID:G14066615)"

misc_feature
66..666
/note="similar to Homo sapiens EST AL522055 (NID:G12785548)"

misc_feature
190..565
/note="similar to Mus musculus EST BG093388

(NID:G12575951) ut70h03.y1"
220..765
/note="similar to Homo sapiens EST BM015896 (NID:G16530250)"
223..798
/note="similar to Mus musculus EST B1854901 (NID:G15995648)"
253
/note="similar to Homo sapiens EST BG251146 (NID:G12760962)"
289..765
/note="similar to Homo sapiens EST AI912697 (NID:G5632552) wt12f12.x1"
305..765
/note="similar to Homo sapiens EST W79609 (NID:G1390880) zd81b07.r1"
335..560
/note="similar to Homo sapiens EST AW938856 (NID:G8114302)"
342..765
/note="similar to Homo sapiens EST AI637897 (NID:G4690131) tt11f10.x1"
364..570
/note="similar to Homo sapiens EST AA428438 (NID:G2110249) zV50a04.r1"
497..836
/note="similar to EST BM676074 (NID:G18985970)"
499..765
/note="similar to Homo sapiens EST BF037254 (NID:G10745479)"
507..765
/note="similar to Homo sapiens EST N56492 (NID:G1199340)"
540..798
/note="similar to Homo sapiens EST AL565079 (NID:G12916097)"
573..798
/note="similar to Homo sapiens EST BE395713 (NID:G9341078)"
607..798
/note="similar to Homo sapiens EST BG149868 (NID:G12661898)"
620..798
/note="similar to Homo sapiens EST BG251146 (NID:G12760962)"
695..798
/note="similar to Homo sapiens EST AU150237 (NID:G11011758)"
738..798
/note="similar to Homo sapiens EST AA284866 (NID:G1927608) zt23a04.s1"
786..794
/note="match to EST AI669845 (NID:G4834619) wb85e08.x1"
811..1179
/note="similar to Homo sapiens EST W46662 (NID:G1331290) zc33d04.r1"
816..1200
/note="similar to Homo sapiens EST AL565079 (NID:G12916097)"
816..1179
/note="similar to Homo sapiens EST BG251146 (NID:G12760962)"
816..1176
/note="similar to Homo sapiens EST BE395713 (NID:G9341078)"
816..1148
/note="similar to Homo sapiens EST AA284866 (NID:G1927608) zt23a04.s1"
816..1148
/note="similar to Homo sapiens EST AU150237 (NID:G11011758)"
816..1103

Query Match

98.0%; Score 4783; DB 9; Length 156442;

QY 2161 AAAAAATCAGCTGTGTGGTAGTATGCTGTAGTCCAGCTATCCAGGAGGCTGAGA 2220
Db |||||
QY 44447 AAAAAATCAGCTGTGTGGTAGTATGCTGTAGTCCAGCTATCCAGGAGGCTGAGA 44506
Db |||||
QY 2221 TGGGAGATCACCTGAGCCCAACAACCTGGAGTCTTGATCATCTACTGAACTGTAGCCTGG 2280
Db |||||
QY 44507 TGGGAGATCACCTGAGCCCAACAACCTGGAGTCTTGATCATCTACTGAACTGTAGCCTGG 44566
Db |||||
QY 2281 GCAACAGAGGATAGTGAGATCCTGTCTCAAAAAAATAATTAATTAAGAGCCAGGAAC 2340
Db |||||
QY 44567 GCAACAGAGGATAGTGAGATCCTGTCTCAAAAAAATAATTAATTAAGAGCCAGGAAC 44626
Db |||||
QY 2341 AAGACTTAGCTCTAACATCTAACATCTAACATCTAACATCTAACATCTAACATCTAAC 2400
Db |||||
QY 44627 AAGACTTAGCTCTAACATCTAACATCTAACATCTAACATCTAACATCTAACATCTAAC 44685
Db |||||
QY 2401 CTGATATTTT - AAAAGTTATAAATAATCTATTAATTCACAATTTGGGGTAAGATAAAGCACT 2459
Db |||||
QY 44686 CTGATATTTTAAAAGTTATAAATAATCTATTAATTCACAATTTGGGGTAAGATAAAGCACT 44745
Db |||||
QY 2460 TGCAGTTTCCAAAGATTTTACAAGTTTACCTCTCATATTTATTTCTTATTTGTGCTATT 2519
Db |||||
QY 44746 TGCAGTTTCCAAAGATTTTACAAGTTTACCTCTCATATTTATTTCTTATTTGTGCTATT 44805
Db |||||
QY 2520 TTAGAGCACCAATATATACTAAATGGAAATGGAAGGGGATTCAGATATTTTCAAAG 2579
Db |||||
QY 44806 TTAGAGCACCAATATATACTAAATGGAAATGGAAGGGGATTCAGATATTTTCAAAG 44865
Db |||||
QY 2580 TGACATATTTGCTGTGGTTAATAATA TGCTCTTTTGTCTGTCAACCAAGGATGGA 2639
Db |||||
QY 44866 TGACATATTTGCTGTGGTTAATAATA TGCTCTTTTGTCTGTCAACCAAGGATGGA 44925
Db |||||
QY 2640 CAGTGATTCAGAACCGTCAAGACGCTAGTGTGA CTTTGGCAGGAATGGGATCCATATA 2699
Db |||||
QY 44926 CAGTGATTCAGAACCGTCAAGACGCTAGTGTGA CTTTGGCAGGAATGGGATCCATATA 44985
Db |||||
QY 2700 AACAGGATTTGCAAAATGTTGCAACCAACAGATGGGAAGAAATTACTGTGGCCTACCAG 2759
Db |||||
QY 44986 AACAGGATTTGCAAAATGTTGCAACCAACAGATGGGAAGAAATTACTGTGGCCTACCAG 45045
Db |||||
QY 2760 GTAAAGAACAGGATGCAAAATAAATAATCAATCTTATTTGAAATGGGATTTTTTAAATTA 2819
Db |||||
QY 45046 GTAAAGAACAGGATGCAAAATAAATAATCAATCTTATTTGAAATGGGATTTTTTAAATTA 45105
Db |||||
QY 2820 AAAACATTTATTTGGAAGCTGTTTATAGGAGTAAAGAGATTTCTGTGACAAAATG 2879
Db |||||
QY 45106 AAAACATTTATTTGGAAGCTGTTTATAGGAGTAAAGAGATTTCTGTGACAAAATG 45165
Db |||||
QY 2880 TGGAAAGCTAAAGATAAGGGAAGAAAGGAGGATTTTATAGTTTCCCAAAATTTATTTTGGT 2939
Db |||||
QY 45166 TGGAAAGCTAAAGATAAGGGAAGAAAGGAGGATTTTATAGTTTCCCAAAATTTATTTTGGT 45225
Db |||||
QY 2940 GAGAGATTTTATTTTGTCTTTTATAGTGAAATATGCGCTGGAAATGAATAAATAGC 2999
Db |||||
QY 45226 --GAGATTTTATTTTGTCTTTTATAGTGAAATATGCGCTTTGGAAATGATAAATAGC 45283
Db |||||
QY 3000 CAGCTTACCAGGATGGACCCACAGAACTTTTGTAGTAAGATGGAGACTGGAAAGGAGAC 3059
Db |||||
QY 45284 CAGCTTACCAGGATGGACCCACAGAACTTTTGTAGTAAGATGGAGACTGGAAAGGAGAC 45343
Db |||||
QY 3060 AAAAGTAAAGGCTCACTATGGAGGATTCAGTGTACAGAAATGAAGCCCAACAAATACCAGATC 3119
Db |||||
QY 45344 AAAAGTAAAGGCTCACTATGGAGGATTCAGTGTACAGAAATGAAGCCCAACAAATACCAGATC 45403
Db |||||
QY 3120 TCAGTGAACAAATACAGAGAACAGCCGATATGCGCTCATGGATGGAGCATCTCAGCTG 3179
Db |||||
QY 45404 TCAGTGAACAAATACAGAGAACAGCCGATATGCGCTCATGGATGGAGCATCTCAGCTG 45463
Db |||||
QY 3180 ATGGGAGAAAAAGGACCATGACCAATTCACACGGCATGTCTTCAGCAGCTATGACAGA 3239
Db |||||
QY 45464 ATGGGAGAAAAAGGACCATGACCAATTCACACGGCATGTCTTCAGCAGCTATGACAGA 45523
Db |||||

QY 3240 GACAATGACGCGCTGGTATGTTGTGGACATCTTTTGTCTCTGCTTTAAAAATCACTAATAT 3299
Db |||||
QY 45524 GACAATGACGCGCTGGTATGTTGTGGACATCTTTTGTCTCTGCTTTAAAAATCACTAATAT 45583
Db |||||
QY 3300 CATTACTCAGNATCATTAACAAATATTTTAAATAGTACCCTCTCTGGGCACTTACTGTG 3359
Db |||||
QY 45584 CATTACTCAGNATCATTAACAAATATTTTAAATAGTACCCTCTCTGGGCACTTACTGTG 45643
Db |||||
QY 3360 AGCCACTGTCTTAAGCTCTTTTATGCACTCACTCGAAAGCAATTTCAACTATAAGGTAGACAT 3419
Db |||||
QY 45644 AGCCACTGTCTTAAGCTCTTTTATGCACTCACTCGAAAGCAATTTCAACTATAAGGTAGACAT 45703
Db |||||
QY 3420 TCTTATTTCTCAATTTTACAGATGAGATTTAGAGAGATTTACGTGATTTGTCCTCAATGTCAAC 3479
Db |||||
QY 45704 TCTTATTTCTCAATTTTACAGATGAGATTTAGAGAGATTTACGTGATTTGTCCTCAATGTCAAC 45763
Db |||||
QY 3480 AACTACCCAGAGATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 3539
Db |||||
QY 45764 AACTACCCAGAGATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 45823
Db |||||
QY 3540 GATAAATACCTATATCTCTATATTCTAAAGTGTGTGTGAAAACTTTTCAATTTTCCA 3599
Db |||||
QY 45824 GATAAATACCTATATCTCTATATTCTAAAGTGTGTGTGAAAACTTTTCAATTTTCCA 45883
Db |||||
QY 3600 GGGTCTCTGATACATAAGGGTTGTAAGAGCTTATTTCCAGATTAAGAGTAAACACACAG 3659
Db |||||
QY 45884 GGGTCTCTGATACATAAGGGTTGTAAGAGCTTATTTCCAGATTAAGAGTAAACACACAG 45943
Db |||||
QY 3660 TCCCTAGATGAGATT - GCCACAAAGCCAGTTATCTCTCTTTCTTGTCTATAGGGCACAGG 3718
Db |||||
QY 45944 TCCCTAGATGAGATTGAGCCACAAAGCCAGTTATCTCTCTTTCTTGTCTATAGGGCACAGG 46003
Db |||||
QY 3719 AGGTCTTTGGTGTATTTAGTGTGACTCTATGTATAGCACCCAAAGAAAGACTTACTGTGCA 3778
Db |||||
QY 46004 AGGTCTTTGGTGTATTTAGTGTGACTCTATGTATAGCACCCAAAGAAAGACTTACTGTGCA 46063
Db |||||
QY 3779 CACGAGTGTAGAGTCTTTTATGGGTAACTGCAAAAGTAACTTGACCAACCGTAGTTCT 3838
Db |||||
QY 46064 CACGAGTGTAGAGTCTTTTATGGGTAACTGCAAAAGTAACTTGACCAACCGTAGTTCT 46123
Db |||||
QY 3839 GTTTCTAATAACGCCAAACACATTTTCTTTCAGGTTTAACTCAGATCCAGATCCAGAAAAACAGTG 3898
Db |||||
QY 46124 GTTTCTAATAACGCCAAACACATTTTCTTTCAGGTTTAACTCAGATCCAGATCCAGAAAAACAGTG 46183
Db |||||
QY 3899 TTTCTAAAGAAAGACGGTGTGGATGGTGTATAATAGATGTCACTGAGCCCAATCCAAACGG 3958
Db |||||
QY 46184 TTTCTAAAGAAAGACGGTGTGGATGGTGTATAATAGATGTCACTGAGCCCAATCCAAACGG 46243
Db |||||
QY 3959 CAGATACCTAGGGGTGACAGTACACCTGGGACATGGCAAGCATGGCACAGATGATGG 4018
Db |||||
QY 46244 CAGATACCTAGGGGTGACAGTACACCTGGGACATGGCAAGCATGGCACAGATGATGG 46303
Db |||||
QY 4019 TGTAGTATGGATGAATTTGGAAGGGGTCTATGGTACTCAATGAGAAAGATGAGTATGAAGAT 4078
Db |||||
QY 46304 TGTAGTATGGATGAATTTGGAAGGGGTCTATGGTACTCAATGAGAAAGATGAGTATGAAGAT 46363
Db |||||
QY 4079 CAGGCCCTTTCTTCCACAGCAATAGTCCCATACTAGATTTTGTCTCTCTGTATGGT 4138
Db |||||
QY 46364 CAGGCCCTTTCTTCCACAGCAATAGTCCCATACTAGATTTTGTCTCTCTGTATGGT 46423
Db |||||
QY 4139 ACAACATTTTGTATCATTTATTTTGAATTTTCTTTCATACATTTATTTCTCTCTTAAA 4198
Db |||||
QY 46424 ACAACATTTTGTATCATTTATTTTGAATTTTCTTTCATACATTTATTTCTCTTAAA 46483
Db |||||
QY 4199 CTCTCAAGCAGAGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAATTTACATTTAAAT 4258
Db |||||
QY 46484 CTCTCAAGCAGAGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAATTTACATTTAAAT 46543
Db |||||
QY 4259 AGCAGATGATTTTCTTTTGTGTTTCTCATTTCTTGTCTCAACCAAGAGTAAACAAAGT 4318
Db |||||
QY 46544 AGCAGATGATTTTCTTTTGTGTTTCTCATTTCTTGTCTCA - CCAAGAGTAAACAAAGT 46602
Db |||||
QY 4319 ATAGTTTTCAGAGAGTGGTGTCTCATATTTTCAGTTCTTAGTTGATTCGAGAAATTTTCAA 4378
Db |||||

```
Db      46603 ATAGTTTTCACAGAGTGGTGTTCATAATTTTCAGTCTTAGTTGATTCGAGAAATTTTCAA 46662
QY      4379 ATAAGGAGAGGGTCTTTATTCCTTGTCTAGGAAACCATGACGGAAAGGAAAACTG 4438
Db      46663 ATAAGGAGAGGGTCTTTATTCCTTGTCTAGGAAACCATGACGGAAAGGAAAACTG 46722
QY      4439 ATGTTTAAAGTCCACTTTTAAACCTATATTTATTTATCTAGGATCTGCAAGAAACT 4498
Db      46723 ATGTTTAAAGTCCACTTTTAAACCTATATTTATTTATCTAGGATCTGCAAGAAACT 46782
QY      4499 TCACAAAGATTTATTAATAACACAGACTCTGTTGCAATAAGTTAAAGTTTCTTGTGTT 4558
Db      46783 TCACAAAGATTTATTAATAACACAGACTCTGTTGCAATAAGTTAAAGTTTCTTGTGTT 46842
QY      4559 TGTAATCCACACATTCATAGTATAGGCTTTCACATTTGAAAGGAGGAGAGCGTTTACA 4618
Db      46843 TGTAATCCACACATTCATAGTATAGGCTTTCACATTTGAAAGGAGGAGAGCGTTTACA 46902
QY      4619 ACCTCAATAGCTAATAAACCGTCTTGAATATTTGAAGATTTAAATCTGACTCTTAGGA 4678
Db      46903 ACCTCAATAGCTAATAAACCGTCTTGAATATTTGAAGATTTAAATCTGACTCTTAGGA 46962
QY      4679 CGGGCACGGTGGCTCAGCACTATAATCCACACACTTTGGAGGCTCAGCGCGGCGTCA 4738
Db      46963 CGGGCACGGTGGCTCAGCACTATAATCCACACACTTTGGAGGCTCAGCGCGGCGTCA 47022
QY      4739 AAGTTCAGGAGTTCAAGACCGCTGACCAATATGTTGAACCCCATCTCTACTAAAAAT 4798
Db      47023 AAGTTCAGGAGTTCAAGACCGCTGACCAATATGTTGAACCCCATCTCTACTAAAAAT 47082
QY      4799 ACAAAATTTAGCAGCGTGGTGGCAGGTCCTGTAGTCCAGCTAGCTGTGAGGTGG 4858
Db      47083 ACAAAATTTAGCAGCGTGGTGGCAGGTCCTGTAGTCCAGCTAGCTGTGAGGTGG 47140
QY      4859 AGATTGCATTGACCAAGATC 4879
Db      47141 AGATTGCATTGACCAAGATC 47161

RESULT 5
AF388026
LOCUS      Homo sapiens fibrinogen, B beta polypeptide (FGB) gene, complete
DEFINITION
ACCESSION AF388026
VERSION    AF388026.1  GI:14423574
KEYWORDS   Location/Qualifiers
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1. (bases 1 to 11956)
AUTHORS   Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
TITLE     Direct Submission
JOURNAL   Submitted (05-JUN-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT   To cite this work please use: SeattleSNPs. NHLBI Program for
            Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.mbt.washington.edu).
FEATURES   Location/Qualifiers
            1..11956
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            repeat_region 1..55
            /rpt_family="L2"
            /rpt_type=dispersed
            variation 40
            /frequency="0.01"
            /replace="G"
            variation 472
```

```
/frequency="0.13"
/replace="A"
675
/frequency="0.01"
/replace="T"
749
/frequency="0.01"
/replace="G"
899
/frequency="0.14"
/replace="T"
1038
/frequency="0.13"
/replace="A"
1437
/frequency="0.13"
/replace="A"
1442
/frequency="0.01"
/replace="T"
1643
/frequency="0.18"
/replace="T"
1744
/frequency="0.13"
/replace="T"
1892..9964
/gene="FGB"
Join(<1892..2013,4689..4880,5370..5553,6474..6701,
7262..7375,8062..8187,8394..8679,9299..9964)
/gene="FGB"
/product="fibrinogen, B beta polypeptide"
Join(1900..2013,4689..4880,5370..5553,6474..6701,
7262..7375,8062..8187,8394..8679,9299..9530)
/gene="FGB"
/codon_start=1
/product="fibrinogen, B beta polypeptide"
/protein_id="AAK62470.1"
/db_xref="GI:14423575"
/translation="MKRMVSWSFHKLTKMKHLLLLCVFLVKSGVNDNERGFSAR
GHRPLDKRREAPSLRPAPPPISSGGYRARPAAKAAQKVKERKAPDAGGCLHAPDL
GVLPTGQLOALQORPIRNSVDLNNVAVSQSSSFQYMLLKDLWKQKQK
QVKNENVEYSSLEKHOLYIDFTVNSIPTNLRLVLSILENLSKIQKLESVSA
QMEYCRTPCTVSCNIPVVSKECEIIRKGETSEMILLQPDSSVKPRVYCDMTEN
GWTIVQRQGSDFGRKWDPIKQGFQGVNATIDGKNYGLDFEGLNIDKISQITR
MGTELLIEMEDWKDKVHAHYGFTVQNEANKYQISVKNYRGTAGNALMDGASQIMG
ENRTMTIHNGMFFSTYDRDNDGLTSDPRKQCKSDGGGWMYNRCHAANPNRGYYWGG
QYTWDMAKHGTDDGVVMNWKGSWYSMRKMSMKIRPFPFQQ"
2240
/gene="FGB"
/frequency="0.13"
/replace="A"
2360
/gene="FGB"
/frequency="0.01"
/replace="G"
2620
/gene="FGB"
/frequency="0.01"
/replace="C"
2623
/gene="FGB"
/frequency="0.02"
/replace="A"
2867
/gene="FGB"
/frequency="0.14"
/replace="A"
2933..2934
/gene="FGB"
/frequency="0.01"
/replace="T"
3395..3471
repeat_region
```

```
/rpt_family="MIR"
3471 /rpt_type=dispersed
    /gene="FGB"
    /frequency="0.03"
    /replace="C"
3582
    /gene="FGB"
    /frequency="0.13"
    /replace="G"
3650. 3780
    /gene="FGB"
    /note="Region not scanned for variation."
3950. .4168
    /rpt_family="MIR"
3958 /rpt_type=dispersed
    /gene="FGB"
    /frequency="0.01"
    /replace="G"
4110
    /gene="FGB"
    /frequency="0.14"
    /replace="I"
4199
    /gene="FGB"
    /frequency="0.14"
    /replace="A"
4219
    /gene="FGB"
    /frequency="0.01"
    /replace="T"
4375
    /gene="FGB"
    /frequency="0.13"
    /replace="C"
4457
    /gene="FGB"
    /frequency="0.02"
    /replace="A"
4512
    /gene="FGB"
    /frequency="0.02"
    /replace="G"
4872
    /gene="FGB"
    /frequency="0.01"
    /replace="I"
4950
    /gene="FGB"
    /frequency="0.13"
    /replace="C"
5616
    /gene="FGB"
    /frequency="0.01"
    /replace="T"
5713
    /gene="FGB"
    /frequency="0.13"
    /replace="I"
5801
    /gene="FGB"
    /frequency="0.01"
    /replace="T"
6491
    /gene="FGB"
    /frequency="0.01"
    /replace="C"
6550
    /gene="FGB"
    /frequency="0.13"
    /replace="T"
6814
```

```
variation
    /gene="FGB"
    /frequency="0.13"
    /replace="G"
6824
    /gene="FGB"
    /frequency="0.13"
    /replace="G"
6943
    /gene="FGB"
    /frequency="0.01"
    /replace="A"
7337
    /gene="FGB"
    /note="Region not scanned for variation."
7445. .7747
    /rpt_family="AluJo"
    /rpt_type=dispersed
7520. .7740
    /gene="FGB"
    /note="Region not scanned for variation."
7742
    /gene="FGB"

Query Match      98.0%; Score 4779.4; DB 9; Length 11956;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4870; Conservative 0; Mismatches 1; Indels 10; Gaps 8;

QY 1 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTTCCAGACCTCCTCTT 60
Db 5431 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTTCCAGACCTCCTCTT 5490
QY 61 CTTCCTTTTCAGTACATGTATTTGCTGAAAGACCTGTGCGAAAAGAGGCGAGCAAGTAA 120
Db 5491 CTTCCTTTTCAGTACATGTATTTGCTGAAAGACCTGTGCGAAAAGAGGCGAGCAAGTAA 5550
QY 121 AAGGTAGATATCCTTGTGCTTTCCATTCGANTTTCAGCTATAAAATTCGAACCGTTAGAC 180
Db 5551 AAGGTAGATATCCTTGTGCTTTCCATTCGANTTTCAGCTATAAAATTCGAACCGTTAGAC 5610
QY 181 TGCCACGAGAATGCATGTTGTGGAAGATTAAACATTTCTGGGTTAGTGAATAGCATTC 240
Db 5611 TGCCACGAGAATGCATGTTGTGGAAGATTAAACATTTCTGGGTTAGTGAATAGCATTC 5670
QY 241 TACGCTTTTGGGCACCTTCCCTCGCACTTCCAGATAAGCACTATTTCAGCTCTTATTTCC 300
Db 5671 TACGCTTTTGGGCACCTTCCCTCGCACTTCCAGATAAGCACTATTTCAGCTCTTATTTCC 5730
QY 301 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAAATTTCTACTATGACTCTTATTTTA 360
Db 5731 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAAATTTCTACTATGACTCTTATTTTA 5790
QY 361 AGTATACAGAAACTTTGTGACTCAGAGATATAATTTTACAGAGTGGAAAAAACCCCTAG 420
Db 5791 AGTATACAGAAACTTTGTGACTCAGAGATATAATTTTACAGAGTGGAAAAAACCCCTAG 5850
QY 421 CATTATAGTTTTAAACATTTTCAGGTTTTTGAATGAGAGATTATTCATTAATATTAATTAAT 480
Db 5851 CATTATAGTTTTAAACATTTTCAGGTTTTTGAATGAGAGATTATTCATTAATATTAATTAAT 5910
QY 481 GTGTTGTGGATAATGACACCTAACCTGTGAATCTTCAGGTGAGATGTTGAGTGTGTTG 540
Db 5911 GTGTTGTGGATAATGACACCTAACCTGTGAATCTTCAGGTGAGATGTTGAGTGTGTTG 5970
QY 541 ACTTGGTGTTCAGGAAACAGCTAGTGTGCGTGGCCCTGGCAGACGGCATCTCAGTGTAGTACA 600
Db 5971 ACTTGGTGTTCAGGAAACAGCTAGTGTGCGTGGCCCTGGCAGACGGCATCTCAGTGTAGTACA 6030
QY 601 TACCCACAGTTGGAAAAATTTTCAAAGAAATCAAAGGAATCATGACATCTTTATAAATTTCA 660
Db 6031 TACCCACAGTTGGAAAAATTTTCAAAGAAATCAAAGGAATCATGACATCTTTATAAATTTCA 6090
QY 661 AGGTTCTGTATCTATGTAATGGAATAAATAAATCAAGCATATCCACTCTGTGAAGAT 720
```



```

Db      54055 ATAAACACTGCGAGTTTCCAAAGATTTTACAGTTTACCTCT-ATATTATTCCTTAT 53997
QY      2510 TGTGCTATTATTAGACCAACAAATATATACTAAATGAA 2548
Db      53996 GGGGGCTATTTTAAAGCCAAATTTTTCCTTACGGGA 53958

RESULT 7
AC009583
LOCUS   Homo sapiens clone RP11-158C21, *** SEQUENCING IN PROGRESS ***, 33
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2769: contig of 2769 bp in length
* 2770 2869: gap of 100 bp
* 2870 6222: contig of 3353 bp in length
* 6223 6322: gap of 100 bp
* 6323 9584: contig of 3262 bp in length
* 9585 9684: gap of 100 bp
* 9685 12108: contig of 2424 bp in length
* 12109 12208: gap of 100 bp
* 12209 16300: contig of 4092 bp in length
* 16301 16400: gap of 100 bp

```

```

* 16401 19929: contig of 3529 bp in length
* 19930 20029: gap of 100 bp
* 20030 23972: contig of 3943 bp in length
* 23973 24072: gap of 100 bp
* 24073 27737: contig of 3665 bp in length
* 27738 27837: gap of 100 bp
* 27838 31345: contig of 3508 bp in length
* 31346 31445: gap of 100 bp
* 31446 36572: contig of 5127 bp in length
* 36573 36672: gap of 100 bp
* 36674 40431: contig of 3759 bp in length
* 40432 40531: gap of 100 bp
* 40532 46767: contig of 6236 bp in length
* 46768 46867: gap of 100 bp
* 46868 53856: contig of 6989 bp in length
* 53857 53956: gap of 100 bp
* 53957 63377: contig of 9421 bp in length
* 63378 63477: gap of 100 bp
* 63478 69622: contig of 6145 bp in length
* 69623 69722: gap of 100 bp
* 69723 75340: contig of 5618 bp in length
* 75341 75440: gap of 100 bp
* 75441 85302: contig of 9862 bp in length
* 85303 85402: gap of 100 bp
* 85403 91207: contig of 5805 bp in length
* 91208 91307: gap of 100 bp
* 91308 98484: contig of 7177 bp in length
* 98485 107735: contig of 9151 bp in length
* 107736 107835: gap of 100 bp
* 107836 119606: contig of 11771 bp in length
* 119607 133091: contig of 13385 bp in length
* 133092 133191: gap of 100 bp
* 133192 145841: contig of 12650 bp in length
* 145842 145941: gap of 100 bp
* 145942 159879: contig of 13938 bp in length
* 159880 159979: gap of 100 bp
* 159980 177241: contig of 17262 bp in length
* 177242 177341: gap of 100 bp
* 177342 194079: contig of 16738 bp in length
* 194080 194179: gap of 100 bp
* 194180 210373: contig of 16194 bp in length
* 210374 210473: gap of 100 bp
* 210474 224591: contig of 14218 bp in length
* 224592 224791: gap of 100 bp
* 224792 240168: contig of 15377 bp in length
* 240169 240268: gap of 100 bp
* 240269 254856: contig of 14588 bp in length
* 254857 254956: gap of 100 bp
* 254957 272278: contig of 17322 bp in length
* 272279 272378: gap of 100 bp
* 272379 302428: contig of 30050 bp in length
* 302429 302528: gap of 100 bp
* 302529 330999: contig of 28471 bp in length.

```

```

FEATURES
            Location/Qualifiers
            1..330999
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="RP11-158C21"
            /clone_lib="RPC1-11 Human Male BAC"

ORIGIN
Query Match      42.9%; Score 2092.4; DB 2; Length 330999;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 2326; Conservative 0; Mismatches 21; Indels 24; Gaps 20;

QY      2528 CCAATATATATACTAAATGGAATGGACAGGGATTGAGATTTTCAAGTGCATTA 2587
Db      46881 CCAAAAGAGACAAATGGAATGGACAGGGATTGAGATTTTCAAGTGCATTA 46940
QY      2588 TTTCGTGTTGGTTAATATATATGCTCTTTTGTCTCAACCAAGGATGGACGTGATT 2647

```

46941	TTTTGCTGTTGGTAAATATATGCTCTTTTGTGTTCTGTCAACCAAGGATGACAGTCAAT	Db	48019	AACACAGTCCCTAGATGAATTGGCCACAAAGGCCAGGTATCTCTCTCTTTCTTGTCTATAGG	48078
2648	CAGAACCGTCAAGAGCGTAGTGTACATTTTCGACAGGAATGGATCCCATATAAAGCAGGA	Qy	3712	GCACAGAGGTCTTTTGGTGTATATTAGTGTGACTCTATGTATAGCACCCCAAGAAAGACTA	3771
47001	CAGAACCGTCAAGAGCGTAGTGTGACATTTTCGACAGGAATGGATCCCATATAAAGCAGGA	Db	48079	GCACAGAGGTCTTTTGGTGTATATTAGTGTGACTCTATGTATAGCACCCCAAGAAAGACTA	48138
2708	TTTGGAAATGTTGCAACCAACACAGATGGGAGAAATTAATCTGTGGCTTACCAGGTAAACGAA	Qy	3772	CTGTGCACACAGGTGTAGCAGTCTTTTATGGGTAAATCTGCAAAAAGTAACTTGACCCCG	3831
47061	TTTGGAAATGTTGCAACCAACACAGATGGGAGAAATTAATCTGTGGCTTACCAGGTAAACGAA	Db	48139	CTGTGCACACAGGTGTAGCAGTCTTTTATGGGTAAATCTGCAAAAAGTAACTTGACCCCG	48198
2768	CAGGCATCCAAATAAATAATCATCTTATTTGAAATGGGATTTTATTTTAAATAAATAAT	Qy	3832	TAGTCTCTGTTT-CTAATAAGCCCAACACATTTTCTTTCAGGTAAACATAGATCCACAGA	3890
47121	CAGGCATCCAAATAAATAATCATCTTATTTGAAATGGGATTTTATTTTAAATAAATAAT	Db	48199	TAGTCTCTGTTTNCATAAGCCCAACACATTTTCTTTCAGGTAAACATAGATCCACAGA	48258
2828	CATTGTTGGAAGCGTGTATTAGGACAGTTAAGAGGAGTTTCTCGACAAATAATGTGGAAGCT	Qy	3891	AAAACAGTCTTCTAAGAGAGACGGTGGTGGTAAATAATAGATGTCATGACGCCAAT	3950
47181	CATTGTTGGAAGCGTGTATTAGGACAGTTAAGAGGAGTTTCTCGACAAATAATGTGGAAGCT	Db	48259	AAAACAGTCTTCTAAGAGAGACGGTGGTGGTAAATAATAGATGTCATGACGCCAAT	48318
2888	AAAGATAAGGGAAGGAGCGAGTTTATAGTTTCCCAAAATTTTATTTTGGTGTAGAGATT	Qy	3951	CCAAACGCGC-AGATACTACTGCGGTGGACAGTACACCTGGGACATGGCAAGCATGGCAC	4009
47241	AAAGATAAGGGAAGGAGCGAGTTTATAGTTTCCCAAAATTTTATTTTGGT- -GAGATT	Db	48319	CCAAACGCGCAGATACTACTGCGGTGGACAGTACACCTGGGACATGGCAAGCATGGCAC	48378
2948	TTATTTTGTCTTTTATAGTGAATATTTGGCTTGGAAATGATATAAATAGCCAGCTTAC	Qy	4010	AGATGATGGTGTAGTATGATGAATTTGGAAGGGGTGATGGTACTCAATGAGGAAGATGAG	4069
47299	TTATTTTGTCTTTTATAGTGAATATTTGGCTTGGAAATGATATAAATAGCCAGCTTAC	Db	48379	AGATGATGGTGTAGTATGATGAATTTGGAAGGGGTGATGGTACTCAATGAGGAAGATGAG	48438
3008	CAGGATGGGACCCACAGAACTTTTCATAGAAATGAGGAGCTGGAAGGAGACAAAGTAAA	Qy	4070	TATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCACATACGTAGATTTTGTCTCTT	4129
47359	CAGGATGGGACCCACAGAACTTTTCATAGAAATGAGGAGCTGGAAGGAGACAAAGTAAA	Db	48439	TATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCACATACGTAGATTTTGTCTCTT	48498
3068	GGCTCACTATGG-AGGATTCACTGACAGAAATGAAGCCCAACAAATPACAGATCTCAGTGA	Qy	4130	CTGTATGTGACAAACATTTTGTGTACATTTTGTATTTTGGAAAT- -TTTCTTTTCATACATTTAT	4188
47419	GGCTCACTATGGNAGATTCACTGACAGAAATGAAGCCCAACAAATPACAGATCTCAGTGA	Db	48499	CTGTATGTGACAAACATTTTGTGTACATTTTGTATTTTGGAAATTTCTTTTCATACATTTAT	48558
3127	ACAAATACAGAGGAACAGCGGTAAATGCTCATGGATGGAGCATCTCAGCTGTGGAG	Qy	4189	TCCTCTAAACCTCTCAAGCAGACGCTGAGTGTGACTTTTGTGAAAAAGTATAGGATAAAT	4248
47479	ACAAATACAGAGGAACAGCGGTAAATGCTCATGGATGGAGCATCTCAGCTGTGGAG	Db	48559	TCCTCTAAACCTCTCAAGCAGACGCTGAGTGTGACTTTTGTGAAAAAGTATAGGATAAAT	48618
3187	AAACAGG-ACCATGACCAATTCACAGCGCATGTTCTTCAGCAGCATGATGACAGACAAAT	Qy	4249	ACATTTAAATAGCACATGATTTCTTTTGTGTTTCTTCTTCTTGTCTACCCCAAGAG	4308
47539	AAACAGG-ACCATGACCAATTCACAGCGCATGTTCTTCAGCAGCATGATGACAGACAAAT	Db	48619	ACATTTAAATAGCACATGATTTCTTTTGTGTTTCTTCTTCTTGTCTAC-CCAAGAG	48677
3246	GACGGTGGTATGTGTGGCAGTCTTTGCTCTGCTTTTAAATACACACTAATATCATTTAC	Qy	4309	TAAACAAAGTATAGTTTGTGACAGATTTGATTTGATTTGATTTGATTTGATTTGATTTG	4368
47599	GACGGTGGTATGTGTGGCAGTCTTTGCTCTGCTTTTAAATACACACTAATATCATTTAC	Db	48678	TAAACAAAGTATAGTTTGTGACAGATTTGATTTGATTTGATTTGATTTGATTTGATTTG	48737
3306	TCAGAAATCAATTAACAAATA- -TTTTTAATAGCTACCACTTCTCT- -GGGACCTTACTGTAGCC	Qy	4369	GAAATTTTCAAAATAAGGAAGAGGGTCTTTTATCTTGTCTGTAGGAAAAACCATGACGAAA	4428
47659	TCAGAAATCAATTAACAAATAATTTTAAATAGCTACCACTTCTCTGGGCGCTTACTGTAGCC	Db	48738	GAAATTTTCAAAATAAGGAAGAGGGTCTTTTATCTTGTCTGTAGGAAAAACCATGACGAAA	48797
3364	ACTGT- -CCTAAGCTCTTTA- -TGCATCAGTCAAGAA- -GCAATTCAGCTATAAGGTAGAC-AT	Qy	4429	GGAAAACTGTATTTTAAAGTCCCACTTTTAAACCTATATTTTATTTATGTAGGATCTGTC	4488
47719	ACTGTCCCTTAAGCTCTTTAATGATCAGTCAAGAAAGCAATTTCAACTAATAGGTAGCAAT	Db	48798	GGAAAACTGTATTTTAAAGTCCCACTTTTAAACCTATATTTTATTTATGTAGGATCTGTC	48857
3420	TCCTATTCTCA- -TTTTACAGATGAGA- -TTTAGAGAGATTACGTGATTTGTCCAA- -TG	Qy	4489	AAAGAAAACTTCCCAAAAAAGTTTATTAATAACAGACTCTGTTTGCATTAAGTTAAATGT	4548
47779	TCCTATTCTCANTTTACAGATGAGNATTTAGAGNAGATTACGTGATTTGTNCCAATGTC	Db	48858	AAAGAAAACTTCCCAAAAAAGTTTATTAATAACAGACTCTGTTTGCATTAAGTTAAATGT	48917
3474	TCACACAACTACCGAGAGATAAATCTAGAAATTTGAGCAGATTACTTCTGAAATATGAG	Qy	4549	TTTCTTGTGTTTGTATCCACACATTTCAATGAGTGGTCTTGTGCACTTGTAAAGGAAGAGA	4608
47839	ACACNAACNTACCGAGAGATAAATCTAGAAATTTGAGCAGATTACTTCTGAAATATGAG	Db	48918	TTTCTTGTGTTTGTATCCACACATTTCAATGAGTGGTCTTGTGCACTTGTAAAGGAAGAGA	48977
3534	CATTTAGATAAATATAC- -TATATCTCTATATCTTAAAGTGTGTGAAAACTTTTCATTTTC	Qy	4609	AGCGTTTCAACCTCAAAATAGCTAAATAAACCGGTCTTGAATATTTTGAAGATTTAAATCT	4668
47899	CATTTAGATAAATATACCTTATATCTCTTAAAGTGTGTGAAAACTTTTCATTTTC	Db	48978	AGCGTTTCAACCTCAAAATAGCTAAATAAACCGGTCTTGAATATTTTGAAGATTTAAATCT	49037
3593	ATTTCCAGGGTCTCTGATACTAAGCGTTGTAAAAAGCTATTATTCCAGTATAAGTAACA	Qy	4669	GACTCTAGACCGGACCGGTGCTCAGCACTATATCCCAACACTTTTGGAGCTGAGGC	4728
47959	ATTTCCAGGGTCTCTGATACTAAGCGTTGTAAAAAGCTATTATCCAGTATAAGTAACA	Db	49038	GACTCTAGACCGGACCGGTGCTCAGCACTATATCCCAACACTTTTGGAGCTGAGGC	49097
3653	AACACAGTCCCTAGATGGATT- -GCCACAAAGGCCAGTTATCTCTCTTTCTTGTCTATAGG	Qy	4729	GGCGGTCTCAAGGTTCAGGATTTCAAGACCGCTGACCAATATGGTGAACCCCATCTC	4788
		Db	49098	GGCGGTCTCAAGGTTCAGGATTTCAAGACCGCTGACCAATATGGTGAACCCCATCTC	49157

QY 4789 TACTAAAAATACAAAAATAGCCAGGCGTGGTGGCAGAGTGCCTGTAGTCCAGCTAGCC 4848
| | | | |
Db 49158 TACTAAAAATACAAAAATAGCCAGGCGTGGTGGCAGAGTGCCTGTGA-GTCCAGCTA-CC 49215
| | | | |
QY 4849 TGTGAGGTGAGATTGCATTGAGCCAAAGATC 4879
| | | | |
Db 49216 TGTGAGGTGAGATTGCATTGAGCCAAAGATC 49246
| | | | |
RESULT 8
AX535001
LOCUS AX535001 4043 bp DNA linear PAT 22-NOV-2002
DEFINITION Sequence 28 from Patent WO02068633.
ACCESSION AX535001
VERSION AX535001.1 GI:25261604
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
AUTHORS Compositions and methods relating to lung specific genes and
TITLE proteins
JOURNAL Patent: WO 02068633-A 28 06-SEP-2002;
Diadexus, Inc. (US)
FEATURES
Location/Qualifiers
Source 1..4043
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 42.7%; Score 2081.2; DB 6; Length 4043;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2083; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2018 CAGGTGTGTGGCTCATACCTGTAAATCCAGACCTTTGGGAGGCCAAAGTGGCTGATAG 2077
| | | | |
Db 1958 CAGGTGTGTGGCTCATACCTGTAAATCCAGACCTTTGGGAGGCCAAAGTGGCTGATAG 2017
| | | | |
QY 2078 CTTGAGCCTTAGAGTTTGAACCTAGCTGGGCAACATATAGAGCCCTAACTCTCAAAA 2137
| | | | |
Db 2018 CTTGAGCCTTAGAGTTTGAACCTAGCTGGGCAACATATAGAGCCCTAACTCTCAAAA 2077
| | | | |
QY 2138 AAAAAAATACCAAAAAAATCAAAAAAATCAGCTGTGGTAGTATGTCCTGTAGT 2197
| | | | |
Db 2078 AAAAAAATACCAAAAAAATCAAAAAAATCAGCTGTGGTAGTATGTCCTGTAGT 2137
| | | | |
QY 2198 CCAGCTATCCAGGAGGTGAGATGGAGATCACTGAGCCCAACCTGGAGTCTTGAT 2257
| | | | |
Db 2138 CCAGCTATCCAGGAGGTGAGATGGAGATCACTGAGCCCAACCTGGAGTCTTGAT 2197
| | | | |
QY 2258 CATGCTACTGAATCTAGCTGGGCAACAGAGGATAGTAGATCTCTCAAAAAAAA 2317
| | | | |
Db 2198 CATGCTACTGAATCTAGCTGGGCAACAGAGGATAGTAGATCTCTCAAAAAAAA 2257
| | | | |
QY 2318 AATTAATTAAGGACAGAAACAGACTTAGCTCTAATCTAATAGCTGACAAAG 2377
| | | | |
Db 2258 AATTAATTAAGGACAGAAACAGACTTAGCTCTAATCTAATAGCTGACAAAG 2317
| | | | |
QY 2378 AGTAATTTGATGTGAATTCACACCTGATATTTAAAGTTTATAAATATCTATAATTCACA 2437
| | | | |
Db 2318 AGTAATTTGATGTGAATTCACACCTGATATTTAAAGTTTATAAATATCTATAATTCACA 2377
| | | | |
QY 2438 ATTTGGGTGAAGATAAGCACTTGAGTTTCCAAAGATTTTCAAGTTTACCTCTCATAT 2497
| | | | |
Db 2378 ATTTGGGTGAAGATAAGCACTTGAGTTTCCAAAGATTTTCAAGTTTACCTCTCATAT 2437
| | | | |
QY 2498 TTATTTCCCTTATGTGCTATTTTATAGACCAAAATATATCTAAATGGAATGACAGG 2557
| | | | |
Db 2438 TTATTTCCCTTATGTGCTATTTTATAGACCAAAATATATCTAAATGGAATGACAGG 2497
| | | | |

QY 2558 GATTCAGATATATATTTTCAAGTGACATATTTTGTGTTGGTTAAATATATGCTCTTTTG 2617
| | | | |
Db 2498 GATTCAGATATATATTTTCAAGTGACATATTTTGTGTTGGTTAAATATATGCTCTTTTG 2557
| | | | |
QY 2618 TTTCTGTCAACCAAGGATGGACAGTGATTCAGAACCGTCAAGACGGTAGTGTGACTTTT 2677
| | | | |
Db 2558 TTTCTGTCAACCAAGGATGGACAGTGATTCAGAACCGTCAAGACGGTAGTGTGACTTTT 2617
| | | | |
QY 2678 GGCAGGAATGGGATCCATATATAAACAAGGATTTGGAAATGTTGCAACCAACAGATGGG 2737
| | | | |
Db 2618 GGCAGGAATGGGATCCATATATAAACAAGGATTTGGAAATGTTGCAACCAACAGATGGG 2677
| | | | |
QY 2738 AAGAATTACTGTGGCTTACCAAGTAACGAGCATGCAAAATATAAATCATTTCTATTG 2797
| | | | |
Db 2678 AAGAATTACTGTGGCTTACCAAGTAACGAGCATGCAAAATATAAATCATTTCTATTG 2737
| | | | |
QY 2798 AATGGGATTTTTTTTAAATTAATAAACAATTCATTTGTTGAAAGCTGTTTAGGCAGTTAA 2857
| | | | |
Db 2738 AATGGGATTTTTTTTAAATTAATAAACAATTCATTTGTTGAAAGCTGTTTAGGCAGTTAA 2797
| | | | |
QY 2858 GAGGAGTTTCTTGACAAAAATGTGAAGCTAAAGATAAGGGAAGAAAGCAGTTTTAGT 2917
| | | | |
Db 2798 GAGGAGTTTCTTGACAAAAATGTGAAGCTAAAGATAAGGGAAGAAAGCAGTTTTAGT 2857
| | | | |
QY 2918 TTCCCAAAATTTTATTTTGGTGAGAGATTTTATTTGTTGTTTCTTTTAGGTGAATTTG 2977
| | | | |
Db 2858 TTCCCAAAATTTTATTTTGGTGAGAGATTTTATTTGTTGTTTCTTTTAGGTGAATTTG 2917
| | | | |
QY 2978 GCTTGGAAATGATAAAATTTAGCCAGCTTACAGAGTGGACCCACAGAACTTTTGATAGA 3037
| | | | |
Db 2918 GCTTGGAAATGATAAAATTTAGCCAGCTTACAGAGTGGACCCACAGAACTTTTGATAGA 2977
| | | | |
QY 3038 AATGAGGACTGGAAAGGAGACAAAGTAAAGCTCACTATGAGGATTCACGTGACAGAA 3097
| | | | |
Db 2978 AATGAGGACTGGAAAGGAGACAAAGTAAAGCTCACTATGAGGATTCACGTGACAGAA 3037
| | | | |
QY 3098 TGAAGCCCAAAATACAGATCTCAGTGAACAAATACAGAGAAACAGCCGGTAATGCCCT 3157
| | | | |
Db 3038 TGAAGCCCAAAATACAGATCTCAGTGAACAAATACAGAGAAACAGCCGGTAATGCCCT 3097
| | | | |
QY 3158 CATGATGGAGCATCTCAGCTGATGGGAGAAAAGGACCATGACCATTCACAAGGCAT 3217
| | | | |
Db 3098 CATGATGGAGCATCTCAGCTGATGGGAGAAAAGGACCATGACCATTCACAAGGCAT 3157
| | | | |
QY 3218 GTTCTTACGACGATGACAGAGCAATGACGGCTGGTATGTGTGGCACTCTTTGCTCT 3277
| | | | |
Db 3158 GTTCTTACGACGATGACAGAGCAATGACGGCTGGTATGTGTGGCACTCTTTGCTCT 3217
| | | | |
QY 3278 GCTTTAAAAATCACACTAATCATCTCAGAAATCATTAACAAATTTTAAATAGTAC 3337
| | | | |
Db 3218 GCTTTAAAAATCACACTAATCATCTCAGAAATCATTAACAAATTTTAAATAGTAC 3277
| | | | |
QY 3338 CACTTCTGGGCACTTACTGTGAGCACTGTCTAAGCTCTTTATGCACTCACTCGAAGC 3397
| | | | |
Db 3278 CACTTCTGGGCACTTACTGTGAGCACTGTCTAAGCTCTTTATGCACTCACTCGAAGC 3337
| | | | |
QY 3398 ATTTCAACTATAAGGTAGACATTTCTATTCTCATTTTACAGATGAGATTTAGAGATTA 3457
| | | | |
Db 3338 ATTTCAACTATAAGGTAGACATTTCTATTCTCATTTTACAGATGAGATTTAGAGATTA 3397
| | | | |
QY 3458 CGTGAATTTGTCCTAATGTCACAACACTACCCAGAGATAAAACTAGAAATTTGAGCACAGTTA 3517
| | | | |
Db 3398 CGTGAATTTGTCCTAATGTCACAACACTACCCAGAGATAAAACTAGAAATTTGAGCACAGTTA 3457
| | | | |
QY 3518 CTTTCTGAATTAATGAGCATTTTAGATAAATCCTATATCTCTATATTTCTAAAGTGTGTG 3577
| | | | |
Db 3458 CTTTCTGAATTAATGAGCATTTTAGATAAATCCTATATCTCTATATTTCTAAAGTGTGTG 3517
| | | | |
QY 3578 AAAAATTTTCAATTTTCCAGGTTCTCTGTACTAAGGGTTGTAAAGCTATTATTC 3637
| | | | |
Db 3518 AAAAATTTTCAATTTTCCAGGTTCTCTGTACTAAGGGTTGTAAAGCTATTATTC 3577
| | | | |

QY 3638 CAGTATTAAGTAACAAACACAGTCCCTAGATGGATGTCACAAAGGCCAGTTATCTCTC 3697
DB 3578 CAGTATTAAGTAACAAACACAGTCCCTAGATGGATGTCACAAAGGCCAGTTATCTCTC 3637
QY 3698 TTTCTTGCTATAGGACACAGAGGTCTTTGGTGTATTTAGTGTGACTCTATGTATAGCAC 3757
DB 3638 TTTCTTGCTATAGGACACAGAGGTCTTTGGTGTATTTAGTGTGACTCTATGTATAGCAC 3697
QY 3758 CAAGGAAGAAGTACTGTGACACAGAGTGTAGTGTGTTTATGCGTAACTCTGCAAAACG 3817
DB 3698 CAAGGAAGAAGTACTGTGACACAGAGTGTAGTGTGTTTATGCGTAACTCTGCAAAACG 3757
QY 3818 TAACTTGACACCGTAGTCTCTGTTTCTAATAACGCCAAACACATTTCTTTCAGGTTAAC 3877
DB 3758 TAACTTGACACCGTAGTCTCTGTTTCTAATAACGCCAAACACATTTCTTTCAGGTTAAC 3817
QY 3878 ATCAGATCCCGAGAAAACAGTGTCTTAAAGAGACAGGTGGTGGATGATATATAGATG 3937
DB 3818 ATCAGATCCCGAGAAAACAGTGTCTTAAAGAGACAGGTGGTGGATGATATATAGATG 3877
QY 3938 TCATGCGACCAATCCAAACGCGAGATCTACTGGGTCGACAGTACACCTGGGACATGCC 3997
DB 3878 TCATGCGACCAATCCAAACGCGAGATCTACTGGGTCGACAGTACACCTGGGACATGCC 3937
QY 3998 AAGCATGGCACAGATGATGGTGTAGTATGGATGAATTTGGAAGGGTCATGTTACTCAAT 4057
DB 3938 AAGCATGGCACAGATGATGGTGTAGTATGGATGAATTTGGAAGGGTCATGTTACTCAAT 3997
QY 4058 GAGGAAGATGATGTAAGATCAGGCCCTTCTTCCACAGCAATAG 4103
DB 3998 GAGGAAGATGATGTAAGATCAGGCCCTTCTTCCACAGCAATAG 4043

RESULT 9
CQ414791/c
LOCUS CQ414791 4967 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 21862 from Patent WO0170979.
ACCESSION CQ414791
VERSION CQ414791.1 GI:41322572
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Lee,J. and Lillie,J.
AUTHORS Genes, compositions, kits, and method for identification,
TITLE assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 21862 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
1..4967
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 19.9%; Score 971.8; DB 6; Length 4967;
Best Local Similarity 99.5%; Pred. No. 4e-190; 2; Indels 3; Gaps 3;
Matches 1006; Conservative 0; Mismatches 0

QY 3869 CAGGTTAAACATCAGATCCCGAGAAAACAGTGTCTTAAAGAGACAGGTGGTGGATGGTGA 3928
DB 2428 CTGTTAAACATCAGATCCCGAGAAAACAGTGTCTTAAAGAGACAGGTGGTGGATGGTGA 2369
QY 3929 TAATAGATGTATGACAGCAATCCAAACCGGAGATCTACTGGGTCGACAGTACACCTG 3988
DB 2368 TAATAGATGTATGACAGCAATCCAAACCGGAGATCTACTGGGTCGACAGTACACCTG 2309
QY 3989 GGACATGGCAAGCATGGCACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4048
DB 2308 GGACATGGCAAGCATGGCACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2249

QY 4049 GTACTCAATGAGGAAGATGATGAAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 4108
DB 2248 GTACTCAATGAGGAAGATGATGAAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 2189
QY 4109 AATAAGTATAGTATTTTCT 4168
DB 2188 AATAAGTATAGTATTTTCT 2129
QY 4169 TTTTCTTCTTATACATATATATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4228
DB 2128 TTTTCTTCTTATACATATATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2069
QY 4229 AAAAAAGTATAGGATAAATATACATTAATAATAGCACATGATTTCTTTTGTCTTCTCAT 4288
DB 2068 AAAAAAGTATAGGATAAATATACATTAATAATAGCACATGATTTCTTTTGTCTTCTCAT 2009
QY 4289 TCTCTCTCTCACCACCAAGTAACAAAGATATAGTTTTCACAGAGTGGTCTTCTCATATAT 4348
DB 2008 TCTCTCTCTCA-CCAGAGATTAACAAAGTATAGTTTTCACAGAGTGGTCTTCTCATATAT 1950
QY 4349 TCAGTCTCTAGTGTGATTCGAGAAATTTTCAAATAAGGAAGAGGGTCTTTTATCTTGTCTG 4408
DB 1949 TCAGTCTCTAGTGTGATTCGAGAAATTTTCAAATAAGGAAGAGGGTCTTTTATCTTGTCTG 1890
QY 4409 TAGGAACCATGAGCGGAAGAAAACATGATTTTAAAGTCCACATTTTAAACTATAT 4468
DB 1889 TAGGAACCATGAGCGGAAGAAAACATGATTTTAAAGTCCACATTTTAAACTATAT 1830
QY 4469 TTATTTATGATAGGATCTCTCAAGAAAACCTTCCAAAAGATTTTATTAATAAACCCAGACT 4528
DB 1829 TTATTTATGATAGGATCTCTCAAGAAAACCTTCCAAAAGATTTTATTAATAAACCCAGACT 1770
QY 4529 CTGTTGCAATAAGTTAATGTTTCTTGTGTTTGTATCCACATTTCAATGATGATGAGGCTT 4588
DB 1769 CTGTTGCAATAAGTTAATGTTTCTTGTGTTTGTATCCACATTTCAATGATGATGAGGCTT 1710
QY 4589 TGCACCTTGTAGGAAGAGAGAGCGTTTCAACCTCAATAGCTAATAAACCGGTCTTGA 4648
DB 1709 TGCACCTTGTAGGAAGAGAGAGCGTTTCAACCTCAATAGCTAATAAACCGGTCTTGA 1650
QY 4649 TATTTGAAGATTTAAATCTGACTCTAGCGGGCAGCGTGGCTCAGCACTATATATCCCA 4708
DB 1649 TATTTGAAGATTTAAATCTGACTCTAGCGGGCAGCGTGGCTCAGCACTATATATCCCA 1590
QY 4709 ACACCTTTGGAGGCTGAGGCGGCGGTCAAGGTTCAGAGTTCAAGACCAGCCTGACCA 4768
DB 1589 ACACCTTTGGAGGCTGAGGCGGCGGTCAAGGTTCAGAGTTCAAGACCAGCCTGACCA 1530
QY 4769 ATATGTTGAACCCCATCTCTACTAATAATACAAAATAGCCAGCGTGGTGGCAGGTG 4828
DB 1529 ATATGTTGAACCCCATCTCTACTAATAATACAAAATAGCCAGCGTGGTGGCAGGTG 1470
QY 4829 CCTGTAGTGTCCAGCTAGCTGTGAGTGGAGATTTGCAATGAGCCAAAGATC 4879
DB 1469 CCTGTGTA-GTCCAGCTA-CCTGTGAGTGGAGATTTGCAATGAGCCAAAGATC 1421

RESULT 10
AX285153/c
LOCUS AX285153 4967 bp DNA linear PAT 20-NOV-2001
DEFINITION Sequence 958 from Patent WO0179556.
ACCESSION AX285153
VERSION AX285153.1 GI:17045841
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Lillie,J., Brown,J.L., Bolt,A. and van Huffel,C.
AUTHORS Novel genes, compositions and methods for the identification,
TITLE assessment, prevention, and therapy of human cancers
JOURNAL Patent: WO 0179556-A 958 25-OCT-2001;

Millennium Predictive Medicine, Inc. (US)

FEATURES

Location/Qualifiers
1. 4967
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

SOURCE

ORIGIN

Query Match 19.9%; Score 971.8; DB 6; Length 4967;

Best Local Similarity 99.5%; Pred. No. 4e-190; Matches 1006; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 3869 CAGGTTAAACATCAGATCCAGAAAACAGTGTCTAAAGAAAGCGGTGGTGGATGGTGA 3928
Db 2428 CTGGTTAAACATCAGATCCAGAAAACAGTGTCTAAAGAAAGCGGTGGTGGATGGTGA 2369
QY 3929 TAATAGATGTCTATGACGCAATCCAAACGGCAGATCTACTCGGGTGGACAGTACACCTG 3988
Db 2368 TAATAGATGTCTATGACGCAATCCAAACGGCAGATCTACTCGGGTGGACAGTACACCTG 2309
QY 3989 GGACATGCCAAGCATGGCAGATGATGGTGTAGTATGGATGAATGGAGGGGTCTATG 4048
Db 2308 GGACATGCCAAGCATGGCAGATGATGGTGTAGTATGGATGAATGGAGGGGTCTATG 2249
QY 4049 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCGCTTCTCCACAGCAATAGTCCCC 4108
Db 2248 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCGCTTCTCCACAGCAATAGTCCCC 2189
QY 4109 AATAGCTAGATTTTGTCTTCTCTATGTGACAAATTTTGTGTACATTTATTTGGAA 4168
Db 2188 AATAGCTAGATTTTGTCTTCTCTATGTGACAAATTTTGTGTACATTTATTTGGAA 2129
QY 4169 TTTTCTTTCATACATATATTTCTCTAAACCTCAAGCAGACGTGAGTGCACCTTTTG 4228
Db 2128 TTTTCTTTCATACATATATTTCTCTAAACCTCAAGCAGACGTGAGTGCACCTTTTG 2069
QY 4229 AAAAAGTATAGGATAAATTCATTAATAATAGCATGATTTCTTTTGTCTTCTCAT 4288
Db 2068 AAAAAGTATAGGATAAATTCATTAATAATAGCATGATTTCTTTTGTCTTCTCAT 2009
QY 4289 TCTTTGTCTACCCAGAAAGTAAACAAAGTATAGTTTTCAGAGATGGTGTTCATAAT 4348
Db 2008 TCTTTGTCTCA-CCAAAGAAAGTAAACAAAGTATAGTTTTCAGAGATGGTGTTCATAAT 1950
QY 4349 TCAGTCTTAGTTGATGCGAGATTTTCAAAATAGGAAGAGGGTCTTTATCCTTGTCG 4408
Db 1949 TCAGTCTTAGTTGATGCGAGATTTTCAAAATAGGAAGAGGGTCTTTATCCTTGTCG 1890
QY 4409 TAGGAAAACCATGACGGAAGGAAAACCTGATGTTTAAAGTCCACTTTTAAACCTATAT 4468
Db 1889 TAGGAAAACCATGACGGAAGGAAAACCTGATGTTTAAAGTCCACTTTTAAACCTATAT 1830
QY 4469 TTATTTATGATAGATCTGTCAAGAAAACCTTCAAAAGATTTATTAATAAACCCAGAT 4528
Db 1829 TTATTTATGATAGATCTGTCAAGAAAACCTTCAAAAGATTTATTAATAAACCCAGAT 1770
QY 4529 CTGTTGCAATAGTTAAATGTTTCTTGTGTTTCTTAATCCACATCAATGAGTTAGGCTT 4588
Db 1769 CTGTTGCAATAGTTAAATGTTTCTTGTGTTTCTTAATCCACATCAATGAGTTAGGCTT 1710
QY 4589 TGCACCTTTGAAGGAGGAGGAGCGTTTCAAACTCAATAGCTAATAAACCCGCTCTTGAA 4648
Db 1709 TGCACCTTTGAAGGAGGAGGAGCGTTTCAAACTCAATAGCTAATAAACCCGCTCTTGAA 1650
QY 4649 TATTTGAAGATTTAAATCTGACTCTAGACGGGCAAGCGTGGCTCACGACTAATAATCCCA 4708
Db 1649 TATTTGAAGATTTAAATCTGACTCTAGACGGGCAAGCGTGGCTCACGACTAATAATCCCA 1590
QY 4709 ACATTTGGGAGGCTGAGCGGGCGGTCTCAAGGTCAGAGTTCAAGACAGCTTGACCA 4768
Db 1589 ACATTTGGGAGGCTGAGCGGGCGGTCTCAAGGTCAGAGTTCAAGACAGCTTGACCA 1530
QY 4769 ATATGTTGAAACCCCATCTCTACTAAATAATACAAAAATAGCCAGCGGTGGTGGCAGGTG 4828

Db 1529 ATATGTTGAAACCCCATCTCTACTAAAAATACAAAAATAGCCAGCGTGGTGGCAGGTG 1470
QY 4829 CCTGTAGTCTCCAGCTAGCTAGCTGTGAGTGGAGATTGCATTGAGCCCAAGATC 4879
Db 1469 CCTGTA-GTCCACAGCTA-CCTGTGAGTGGAGATTGCATTGAGCCCAAGATC 1421

RESULT 11

AC140369/c

LOCUS

DEFINITION

AC140369 Mus musculus chromosome UNK clone RP23-200L3, WORKING DRAFT

SEQUENCE, 9 unordered pieces.

AC140369

AC140369.1 GI:28475627

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 223534) Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

McPherson, J.D. and Waterston, R.H.

2 (bases 1 to 223534) The sequence of Mus musculus clone

Unpublished

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

Center project name: M.BA0200L03

Sequencing vector: M13; 0%

Chemistry: Dye-primer ET; 0% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 218876 bases at least Q40

Consensus quality: 219375 bases at least Q30

Consensus quality: 219567 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently

* consists of 9 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 4212: contig of 4212 bp in length

* 4213 4312: gap of unknown length

* 16828 16928: contig of 12516 bp in length

* 16929 29505: contig of unknown length

* 29506 29605: contig of 12577 bp in length

* 29606 42552: gap of unknown length

* 42553 42552: contig of 12947 bp in length

* 42553 59543: gap of unknown length

* 59544 90619: contig of 16891 bp in length

* 90620 90719: contig of 30976 bp in length

* 90720 114137: contig of 23418 bp in length

* 114138 114237: gap of unknown length

* 114238 159401: contig of 45164 bp in length

* 159402 159501: gap of unknown length

* 159502 223534: contig of 64033 bp in length.

1688 TCTAATATTGTTATATTTTAAAGGCTCTATATTAATACACACACTCCTTCTAGTAACCTTAGT 1747
167491 TCCAACTCTAGAAACATTAATGTTTCAGATGTTTGTGATATAAATGGTCTCTTACCAGTGG 167432
1748 TAAATGTTATTTTAAAGAAATGGTGACTATAATAACAAAGTAATATATGTATATAAACCCTGAA 1807
167431 AACAAATCAGTTATGTTTCAAGAGGAATAGCTACA---GGATCTGTTACTGTCCCTGAACA 167375
1808 CATATATGTTGTTCTTACATTTGAGAAATGTCAGAAATTAATATCAGGAAAGAGAGGTGAAACAT 1867
167374 TAAATGTTCTATCATATTTCTGAGAGTGTGAGGAGATCAATTAGGAAGGAGGTGAGACAT 167315
1868 CTGAATATGTTCTCATTTCAACTGACAGTTCTGTCAAAACCGTATAGAGTATAGTGTGACA 1927
167314 CCGAAATGTTATCTATCCAGCTGACCTCCATCAAGCCGTACAGAGTATATCTGTGACA 167255
1928 TGAATA CAGAAATGAGGTAAGCTTTTCGACAGTTGTTGACCTGTTGATCTGTAATATT 1987
167254 TGAATA CAGAAATGAGGTAAGTGTGATGCTCAGTGGCTGACCTGTCATTCTGTAAGCATC 167195
1988 TGGATACCGTAAATGCGCAGGAAACAGGCCAGGTGTGGTGGCTCATACCTGTAATTTCCA 2047
167194 CTGGGTGCTGGGGAACAAGG----- 167174
2048 GCACCTTGGGAGCGCAAGTGGGCTGATAGCTTTGAGCTAGGCTTTGAAACTAGCCTGG 2107
167173 ----- 167174
2108 GCAACATATGAGACCTTAACCTTACAAAAAATAAATAAATAAATAAATAAATAAATAA 2167
167173 ----- 167174
2168 TCAGCTGTGTTGGTAGTAGTGTGCTGTAGTCCAGCTATCCAGGAGCTGAGATGGGAGA 2227
167173 ----- 167174
2228 TCACCTGAGCCCAACCTGGAGTCTTGATCATGCTGTAACCTGTAAGCTGTGGCAACAG 2287
167173 ----- 167174
2288 AGGATAGTGAGATCTCTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 2347
167173 ----- 167174
2348 AGCTCTAACATCTAACATAGCTGACAAAGAGTAAATTTGATGTGGAAATTCACCTGATAT 2407
167173 -----TTACGCTCCAAATTCAGATGACAAAGGAACCACTT----- 167141
2408 TTAAAGTTTATAAAATATCTATAATTCACAATTTGGGGTAAGATAAAGCACTTGCAGTTT 2467
167140 -----GATTTGAAGAACTCGATGCGATGCAATGAAGGTGTCAATGACGACCTTGAGATT 167087
2468 CCAAGATTTTCAAGTTTACCTCTCATATTTATTTCTTATTTGTTCTATTTT-----TAG 2523
167086 CCAATACTTTAAGAGTCTCTCTTATGCTTATTTCTTATCACAATCCCTTTGGAGAG 167027
2524 AGCACCATAATATACTAAATGGAATGGACAGGGATTCAGATATATTTTCAAAGTGAC 2583
167026 GGACCATAGGAGGCATAAGATAAACAATTTGGCAATAAGTTTCAGACTTTTCAGAGTAAT 166967
2584 ATTTATTTGTTGTTTAAATATATGCTTTTGTCTTCTCAACCAAGAGATGGACACT 2643
166966 GGTATTTAG-----TGATTTCTTAGGCTAAATTTTATTTCTCTCAACTAAAGGATGGACGT 166911
2644 GATTTCAAGAACCGTCAAGACCGTAGTGTGACCTTTGGCAGGAATGGGATCCATATAAACA 2703
166910 CATACAGAACCGTCAAGATGGACGCTCGACCTTTGGCAGGAATGGGATCCATATAAACA 166851
2704 GGGATTTGGAATGTTTGCACCAACACAGATGGGAAGAAATTTACTGTGGCTACCGGTA 2763
166850 AGGATTCGGAATATTTGCCCAACCAATGAAGACGGAAGAAGTACTGTGGCTTCCAGGTAA 166791

2764 CGAACAGCGCATGCAAAAAATAAATCATTTATTTGAAATGGGATTTTATTAATTAATAAAA 2823
166790 AGACTGCGAGTGAAAAAATCAATCCTTCTATTTTAAAGAGTTTCTTCTCAGTCAAAA 166731
2824 CATTCATTGTTGGAAGCGCTGTTTTAGG----- 2850
166730 CATCAAGGATGGGATCTCTTTTTAGAAATTTTAGGATGCCAGGAAGATTTTAACTACAC 166671
2851 -----CAGTTAAGAGGAGTTTCTCGACAAAAATCTGACAAAAATCTGGAAGCT 2887
166670 ACATGCACTACACACATATGACACATAAATAAATGTAATGTTTAAATGTCGAGCT 166611
2888 AAGATAAGGGAAGAAAGCAGCTTTTATGTTTCCCAAAATTTTATTTTGGTGAGAGATT 2947
166610 AAGGATCTGGGAGAA--GCTGTGCCCCGTTGCCCTCGGCTTTACTTTTGTGATGGCTT 166553
2948 TTATTTTGTGTTTTTCTTTTAGTGAATATTTGGTTGGAAATGATAAATTTAGCAGCTTAC 3007
166552 ATA---TCCATTCTCTGTAGTGAATATTTGGTTGGCAATGATAAGATTAGCCAGCTAC 166496
3008 CAGGATGGGACCCACAGAACTTTTGTAGTAATGAGGAGCTGGAAGGAGCAAAAGTAAA 3067
166495 CAGGATGGGACCCACAGAACTTTCTCATTTGAATGGAGGACTGGAAGGAGACANGTGA 166436
3068 GGCTCACTATGAGGATTCACGTACAGAAATCAAGCAAAATACCAAGATCTCAGTGA 3127
166435 GGCATTTATGAGGCTTCACGTTACAGAAACGAGGCGCAATATACCAAGTCTCAGTGA 166376
3128 CAAATACAGAGAAACAGCGGTAATGCTCATGATGAGGATCTCAGCTGATGGAGA 3187
166375 CAAATACAAAGGACGCGTGGCAACGCTCATGACGCGGCTCTCAACTGTGGGGA 166316
3188 AAGAGGACCATGACATTCACACGCGCATGTTCTTACGACGTATGACAGACATGA 3247
166315 GAACAGAACCATGACCATTCACACGCGCATGTTCTTACAGCGTACGACGAGGACAA 166256
3248 CGGCTGGTATGTTGGGCACTTTTGTCTCTG--CTTTAAATAACACACTAATATCACT 3306
166255 CGGCTGGTAAAGCAACCAAGGCTTAGTCTCTGTTGATGATGACCCCAAGANTGGATTGTC 166196
3307 CAGAAATCAATA-----CAATATTTTAAATAGTACCACTTCTTGGGCACTTA----- 3354
166195 TAGCACTGTGGGGTACCGAATGTTCTTAAACAACTGCCAGACTTGGACACTTCAAGTGT 166136
3355 --CTGTACGACCTGCTCTAGCTCTTATGATCACTCGAAGCATTTCAACTATAAGG 3412
166135 ATCAGGACGCCCCATCCGAACTCTTACGTTATCAAGGGAAGCATCCAGCTAAGA 166076
3413 TAGACATTTCTTATTTCTCATTTTACAGATGAGATTTAGAGAGATTACGTATTTGCCAAT 3472
166075 CAGACATGGTTAGTTTCTCTTTGAGATGAGTTTGTAGATACGATTTTACCC----- 166023
3473 GTACACAACTACCCAGAGATAAATACTAGAAATTTGAGCAGCTTACTTTCTGAATAAGA 3532
166022 --TAAGACAGCTACCCAAAGATAGACCTGTATCTATAGCACAAAC-----TCTGAGTTTTGT 165969
3533 GCATTTAGATAAATACCTATATCTCTATTTCTTAAAGTGTGTGAAACCTTTCAATTTTC 3592
165968 GTCTCTATATACATCTGTGTTTTTCACTCAACGCTATCTACGGAAGCCATCATCCA 165909
3593 ATTTCCAGGTTCTCTGATACCTAAGGTTGTAAGCTATTTTCCAGTATATAAGTAAACA 3652
165908 TATCTCAGGCT-----CTCATCTACGATATCATGAAG 165872
3653 AACACAGTCCCTTAGATGGATTG--CCAAAAAGGCCAGTTATCTCTTTTCTGCTATAGG 3711
165871 AATTAGTCCCTTCAATGATCAGATCAGATAGACCCAGCTGTTTCCCTTTCTGCTATGG 165812
3712 GCACAGGAGGCTTTTGGTGTATTTAGTGTGACTATATGATATAGCAACCAAGGAAGACTA 3771
165811 GCAGAGAAGAACTTGGCTGTATTTG---GGCTTTACATACAGGATTCAGGGAAGAGGT 165755
3772 CTGTGCACAGAGTGTAGCAGTCTTTTATGGGTAATCTGCAAAACGTAACCTTGACCCCG 3831

```

165754 CCACACAGTGGAGTACAATACCTTTTCATCGGGTACTCTCCAAAGGACTGAATTTGATGG 165695
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
3832 TAGTCTCTGTTCTTAATACGCGCAACA--CAATTTCTTTTCAGGTTAAACATCAGATCCAG 3889
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
165694 TGGCAGCTGTTATGAGTACAATACTACATCATCTCTCTTTTCAGGGTAACTACGAGTCCAG 165635
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
3890 AAAACAGTGTCTTAAAGAGAGCGGTGGTGGATGGTGTATTAATAGATCTCATGCGAGCCAA 3949
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
165634 AAAACAGTGTCTTAAAGAGAGCGGTGGTGGATGGTGTATTAATAGATCTCATGCGAGCCAA 165575
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
3950 TCCAAACGCGAGATACCTACTCTGGGTGGAGCATGACCTGGGACATGGCAAGCATGGGAC 4009
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
165574 TCCAAATGGCAGATACCTACTCTGGGTGGAGCATGACCTGGGACATGGCAAGCATGGGAC 165515
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
4010 AGATGATGCTGTAGTATGATGAGTAATGGAAGGGTCACTGCTACTCAATGAGAGATGAG 4069
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
165514 AGATGATGCTGTAGTATGATGAGTAATGGAAGGGTCACTGCTACTCAATGAGAGATGAG 165455
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
4070 TATGAAGATCAGGCGCTCTTCCACACAGCAATAGTCCCAATACGATAGATTTTGCCTCTT 4129
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
165454 TATGAAGATCAGGCGCTCTTCCACACAGCAATAGTCCCAATACGATAGATTTTGCCTCTT 165395
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
4130 CTGATGTCGACAAATTTTGTATGATGATTTTGGAAATTTCTTTTCATACATTATTT 4189
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
165394 TTCTATGT-CTGACATTTTTCATATTATGTTATTTGGAATTTCTTTTCATACATTATA-T 165337
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
4190 CCTCTAAACCTCAAGCAGAGCTGAGTGCATCTTTTGAAGAAAGTATAGATAAATTA 4249
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
165336 CACTTAAAGCTCAAAACAAATGTTGAGTGTCTCTTTTGGAAACAAATAGATAAAGAC 165277
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
4250 CATTAAATAGCAGATGATTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTT 4292
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |
165276 ATTGAAGAACAGCACCCTAGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 165234
|   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |   ||| |

RESULT 12
AC138394/c 227600 bp DNA linear ROD 13-APR-2004
LOCUS Mus musculus chromosome 3, clone RP23-333D8, complete sequence.
DEFINITION AC138394
ACCESSION AC138394
VERSION AC138394.7 GI:46359988
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 227600)
MUS musculus chromosome 3, clone RP23-333D8, complete sequence.
TITLE AC138394
JOURNAL Unpublished
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Hafez,N., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Lander,R., Levine,R.,
Lindblad-Toh,K., Liu,X., Maclean,C., Macdonald,P., Major,J.,
Mathews,C., McCarthy,M., Melgrim,J., Meneus,L., Mihoval,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rogov,P.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodor,J., Topham,K., Travers,M.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (28-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
1 (bases 1 to 227600)
MUS musculus chromosome 3, clone RP23-333D8
TITLE AC138394
JOURNAL Unpublished
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Hafez,N., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Lander,R., Levine,R.,
Lindblad-Toh,K., Liu,X., Maclean,C., Macdonald,P., Major,J.,
Mathews,C., McCarthy,M., Melgrim,J., Meneus,L., Mihoval,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rogov,P.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodor,J., Topham,K., Travers,M.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (28-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

REFERENCE
AUTHORS

3 (bases 1 to 227600)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Lander,R.,
Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Melgrim,J., Meneus,L., Mihoval,T., Meneus,L., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S.,
Theodor,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 227600)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lander,E.,
Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Melgrim,J., Meneus,L., Mihoval,T., Meneus,L.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,C., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodor,J., Topham,K.,
Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

Direct Submission

Submitted (13-APR-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 13, 2004 this sequence version replaced gi:45504259.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIDR

Web site: <http://www-seq.wi.mit.edu>Contact: sequence.submissions@broad.mit.edu

----- Project Information

Center project name: L28917

Center clone name: 333_D_8

FEATURES

source

Location/Qualifiers

1. 227600

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="3"

/map="3"

/clone="RP23-333D8"

/clone_lib="RPCI-23 Female Mouse BAC"

1..10919

/note="wgs_end_extension"

misc_feature

```

repeat_region 1. .267
/rpt_family="L1_MM"
repeat_region 269. .319
/rpt_family="A-rich"
repeat_region 349. .464
/rpt_family="PB1D7"
repeat_region 515. .658
/rpt_family="URR1A"
repeat_region complement(753. .832)
/rpt_family="BC1_MM"
repeat_region 1057. .1088
/rpt_family="CAGA)n"
repeat_region 1091. .1147
/rpt_family="GA-rich"
repeat_region complement(2211. .2449)
/rpt_family="B4A"
repeat_region 4164. .4310
/rpt_family="B1_MM"
repeat_region 4600. .4907
/rpt_family="ORR1A3"
repeat_region 5091. .5119
/rpt_family="(TTTG)n"
repeat_region complement(5120. .5266)
/rpt_family="B1_MM"
repeat_region complement(5369. .5515)
/rpt_family="MTE"
repeat_region 5627. .5857
/rpt_family="RLTR14"
repeat_region 5858. .6039
/rpt_family="(CTTG)n"
repeat_region 6040. .6164
/rpt_family="RLTR14"
repeat_region complement(6559. .7236)
/rpt_family="L1_MM"
repeat_region complement(7237. .7373)
/rpt_family="B1_MM"
repeat_region complement(7374. .7936)
/rpt_family="L1_MM"
repeat_region 7934. .9190
/rpt_family="L1"
repeat_region 10167. .10291
/rpt_family="Tigger6a"
repeat_region 10311. .10728
/rpt_family="Tigger6a"
misc_feature 10920. .10925
/note="clone_boundary"
clone_end:SP6
site:ECORI"
repeat_region 11324. .11522
/rpt_family="B3"
repeat_region 11748. .12040
/rpt_family="B4"
repeat_region 12268. .12302
/rpt_family="(TA)n"
repeat_region complement(12965. .13038)
/rpt_family="ID5"
repeat_region complement(13425. .14334)
/rpt_family="Lx9"
repeat_region complement(14508. .14803)
repeat_region 15320. .15486
/rpt_family="LIM4"
repeat_region 15558. .15607
/rpt_family="RSINE1"
repeat_region complement(15710. .15899)
/rpt_family="LIME3A"
repeat_region 17549. .17775
/rpt_family="B3"
repeat_region 17969. .18172
/rpt_family="B3"
repeat_region 19415. .19531
/rpt_family="B1_MM"
clone_end:SP6"
1. .267
/rpt_family="(TA)n"
complement(20163. .20472)
/rpt_family="L2"
20521. .20558
/rpt_family="(TG)n"
complement(20687. .21055)
/rpt_family="MT2A"
complement(21237. .21669)
/rpt_family="L2"
22687. .22713
/rpt_family="(TTTG)n"
complement(22716. .22838)
/rpt_family="B1_MM"
23789. .24147
/rpt_family="ORR1D"
complement(24968. .25107)
/rpt_family="B1_MM"
25422. .25466
Query Match 14.3%; Score 696.2; DB 10; Length 227600;
Best Local Similarity 56.1%; Pred. No. 1.9e-133;
Matches 2480; Conservative 0; Mismatches 1298; Indels 645; Gaps 36;
QY 1 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCCGAGACTCTCTCTT 60
DB 121355 TCAGAGTAGTAGTGTGCTGAGTTAATAACATACATACAGTCTGTTCCGATACCTCTCTAG 121296
QY 61 CTTCTTTTCAGTACATGATTTTGTGGAAGACCTGTGGCAAGAGGCGAGAGCAAGTAA 120
DB 121295 TCACCTTTTCAGTACCTGACTGTCTAAAGGACATGTGGAAGCAAGCTCAAGTTA 121236
QY 121 RAGTAGATATCCT-----TGTCCTTCCATTCGATTTTCAGTATAAATTTG 168
DB 121235 AAGGTAGGCTTCTCAGGGGGGGGGGGGGTCTTACCTAACTCTTTTATAAATTTG 121176
QY 169 GAACGGTTAGACTGCACGAGAAATGCATGTTGTGAGAAAGTAA----- 212
DB 121175 AAACGTGTTAACTGTGTATACACAGTCTCTCCCATTTAAACCATCAGGGTTAACCGAC 121116
QY 213 -----ACATTTCTGGGTTAGTGAATAGCAATCATAGCTTTTGGGCACCTTCCCTG 264
DB 121115 TTGCTCTGATTTCTGCTAGTCCCTGCTCTCTCTGCTGCTCTCCCTCAGGCACACTCATG 121056
QY 265 GAACCTGCCAGATAGCACTATTGAGCTCTTATTCAGCTCTGATCAGATCAGCAAGTGTGAT 324
DB 121055 CAGTTTGGCCAGACAAG-ACCTCTCAAGTGTGTAGCCCTTGTGCGCACTTTCACACTACA 120997
QY 325 TTCTCTATGAAAAATTTCTACTATGACTCTCTTATTTTAAAGTATACAAGAAACTTTGTACTCA 384
DB 120996 TTCTCTTGGAGAAATGCTCACTCGGGCTTCTTACTGTGCTATTAACAATTTGTGACTCA 120937
QY 385 GAAGATAATATTTACAGAGTGGAAAAAACCCTAGCATTTTATAGTTTAAACATTGAGG 444
DB 120936 GAATATAATATGCATAG-----GATAAAACCTTAGTATTTGTAAGGTACTCAAT----- 120888
QY 445 TTTTGAATGAGAGAGCTTATCCATAATATATTCAAATGTGTGTGATATGACACCTTAC 504
DB 120887 -----ACCTGAGAGTACTGAAATATGAGAACCATCAGTGT-----GATATAATCTGTAA 120838
QY 505 CTGTGATCTTTGAGGTCAAGAAATGTTGAGTGTGTTGAGCTTGTGCTCAGGAAACAGCTAG 564
DB 120837 CAGTGACATCTGGCATACAAATCTTTGAGGGCTACGGGCTGTGTTGGTGGAAATCAGTTGG 120778
QY 565 TGCCTGAGCTGGCAGAGGCATCTCAGTGAGTAGCATACCCACAGTTGGAAATTTTCAA 624
DB 120777 GGTGTCA-CATGGCAGAGGCTTTTCAGAGAGCAGCAGACCCCAAGGAGTACACAGTTA 120719
QY 625 AGAATCAAGGATCATGACA--TCTTATAAATTTCAAGGTTCTGCTATATCTTATGTA 682
DB 120718 GGCATCAAAATGAATTCAGAAGTGTCTTTGTAATCCTTAAGGCTTTGCAAGTACCCATGGG 120659
QY 683 AATGATAAATAAATCAAGCATATCCACTCTGTAGATTTGAACTTCTCAGATGGAAGACC 742

```

Db	120658	TGTG-----CGAGTGCCAACTGCCTCTTCTCTGCAATGCTATGAGATGGAGCCT	120610	Db	119644	AACAATCAGTTATGTTCAAGAAGGAATAGCTACA---GGATCTGTTACTGTCCCTCGAACA	119588
Qy	743	CCAATACCTGCTTCTCTCTCTTCTCCCTCACCAAGAAATAACAACCTATTTCATTTATT	802	Qy	1808	CATAATGTTGCTTTACATTTGAGAGAAATTTATCAGGAAAGAGAGGTGAAACAT	1867
Db	120609	TCAGTACTGCTCCC-----AAAGTAAGACATGACTTGTGTGCA	120572	Db	119587	TAAATGTCATCATATCTTTCAGAGTGTGAGGAGATCATTTAGGAAGGAGGTGAGACAT	119528
Qy	803	ACTGACACAAATCTTTAGCGTATACCTATGGTAATACTAGTAAATTTAGTGGTGGTATGATTTA	862	Qy	1868	CTGAAATGTATCTCATTTCAACCTGACAGTTCTGTCAACCGTATAGAGTATACCTGTGACA	1927
Db	120571	CCTAGGACAAATCACTGGAGTAGTAAGTAGTGAAGTCACTGATGCAATTTATACATGCCAT	120512	Db	119527	CGAAATGTATCTCATTCAGCTGACACCTCCATCAAGCCGTACAGAGTATACCTGTGACA	119468
Qy	863	TGTTAAATTTGATATGTCATGGCGCAAAATCATTTCCACTAAATATGACTATATATCATPAA	922	Qy	1928	TGAATACAGAAATGGAGGTAAAGCTTTCCAGAGTGTGTTGACCTGTGATCTGTTAATTTAT	1987
Db	120511	GAAGGAAGAAACCAACAAACAAACAAATATGCTCTCTAAAGTTTCAGCA---GTCACGA	120455	Db	119467	TGAATAACAGAAACGAGGTAAAGTATGCTCAGTGGCTGACCTGTCATCTCTGTAAGCATC	119408
Qy	923	CTGCTTGTGTGATAGCTCAGTGTGTTAATAGTTTATTTCTCAGAAATCAAAATTTGAT--AG	980	Qy	1988	TGGATACCGTAAAAATGCCAGGAAACAAGCCAGGTGTGGTCTCATACCTGTATTTCCA	2047
Db	120454	TTGCTTAGTGCCACCCAAAGTATTTAAGTTTATTTCTCAGAGATCAAAATCATATGAT	120395	Db	119407	CTGGGTGCTGGGAAACAAGG-----	119387
Qy	981	TTAATACATTTAGTTTTATGAGGCAAAATGCTAACTATTTCTACATAATTTCATTTTTC	1040	Qy	2048	GCACCTTGGGAGGCCAAAGTGGGCTGATAGCTTGAGCCTAGAGTTTGAACCTAGCCTGG	2107
Db	120394	GTAGTTAGATATCTTTTATAGGCAAAATATGCTGACCACTCCTGTATCATTTTGGCTTTC	120335	Db	119386	-----	119387
Qy	1041	CAGATAATGAAATAGTCAATGAGTACTCCTCAGAACTGGAAAGACCAATTTATATA	1100	Qy	2108	GCAACATAATCAGACCCCTAACTCTACAAAAAATAAATAAATAAATAAATAAATAAATAA	2167
Db	120334	CAGAAATGAAATGTCATAATGAGTACTCCTCCATCCTCGAAGATCAGAGATTATATA	120275	Db	119386	-----	119387
Qy	1101	TAGATGAGACTGTGAATAGCAATATCCCAACTPAACCTTCGTGCTCGTTCAATCCCTGG	1160	Qy	2168	TCAGCTGTGTGGTAGTATGTGCTGTAGTCCAGCTATCCAGGAGGCTGAGATGGAGA	2227
Db	120274	TGATGAGACTGTGAATGATATATCCCTTATCTTCGTGCTCGCTCAATCCCTGG	120215	Db	119386	-----	119387
Qy	1161	AAAACTGAGAGCAAAATACAAAAAGTTAGATCTGATCTCTCAGCTCAAAATGGAATATT	1220	Qy	2228	TCACCTGAGCCCAACACTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGGGCAACAG	2287
Db	120214	AGGACCTGAGAAGCAAAATACAAAAATTTAGAGTCTGACATCTCAGCTCAAAATGGAATACT	120155	Db	119386	-----	119387
Qy	1221	GTGCAACCCCATGCACTGTGCTGCTGCAATATTCCTGTGCTGCTGGCAAGTAAGTACTGAT	1280	Qy	2288	AGGATAGTGAATCCTGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA	2347
Db	120154	GCCGCACTCCATGCACTGTGCTGCTGCAATATTCAGTGGTGTCTGGGAAAGGTAAGTACT	120095	Db	119386	-----	119387
Qy	1281	TCATAAACATATTTTATGAGAGTTCAGAGAACTCCACACCAAAATAAAGAGAACAC	1340	Qy	2348	AGCTCTAATCTAATCATAGCTGCAAGAGGAGTAAATTTGATGTGGAATTTCAACCTGATAT	2407
Db	120094	TAGTGAATACATTTCCATAGACTTCTGAGAGTCTAACTCTCTTTAAACATAGGAGCTAA	120035	Db	119386	-----TTACGTCTCAATTCAGATGACAGGAACCACTT-----	119354
Qy	1341	AACAAACAAACAAATGCTAAGTGGATTTTCCCAACAGATCATPAATGACATTACAGTACAT	1400	Qy	2408	TTAAAGTTTAAATATCTATATTTCAAAATTTGGGTGAGTAAAGTAAAGCACTTGACGTTT	2467
Db	120034	AAAGTGAGCAGAGTAGCT-----TTTCCCCAGCAGAGATGTAATAATTTCTAGTATAT	119982	Db	119353	-----GATTTGAAGAACTCGATGTCATGATGAAGGTGTCAATGACAGCTTTGAGATT	119300
Qy	1401	CATAAAATATCTTAGCCAGTGTGTTTGGACTGGCTGGTGCATTTGCTGGTTTGA	1460	Qy	2468	CCAAAGATTTTACAGTTTACCTCTCATATTTATTTCTCTTATTTGTTCTATTT---TAG	2523
Db	119981	CATGTTATATCTCGGGCATGTCTATTTCTGCTTGAATACTCAACCTCCACCCCGTCTGCT	119922	Db	119299	CCAAATCTTTAAGAGTCTCCCTTTATGCTTTATTTCTTTATTCATCATCCCTTGGAGAG	119240
Qy	1461	TGAGCAGGATGGGCACAC-----GGTAGTCCAGGGGTGGCTGATGTGTCATCTGC	1512	Qy	2524	AGCACCATAATATATACTAAATGGAATGGACAGGGGATTCAGATATTTTTCAAAAGTGAC	2583
Db	119921	TGAGCGGATTAATATATATGAGGTGGGCATTTTCAAGGGTCTTAATGTGTATGGGT	119862	Db	119239	GGACCATAGGAGGCATAAGATAAAACAAATTTGGCAATAAGTTTCAGAGTAAAT	119180
Qy	1513	GPACTGGCTTGAACAGATGGCAGAAC-----ACAGATAGATGTAGAAAGTTTCTCCATTT	1567	Qy	2584	ATTATTTGCTGTGTTTAAATATATGCTCTTTTGTCTTTGTTCTCTGTCACCAAGGATGACAGT	2643
Db	119861	GTGCTGTGTGAACAGATGGCAAGGCCATGGAAGGAAAGATGAGAGAGTTTCTTCATCC	119802	Db	119179	GGTATTAG-----TGATTACCTTAGGCTAAATTTATTTCTCTCACTAAAGGATGACAGT	119124
Qy	1568	TGTTGTGCTGGAGCTCATGATATTTCCAGACACAAAGGTGGAGAGCTTTGTTTC	1627	Qy	2644	GATTTCAGAACCGTCAAGACGGTGTGTTGACTTTTGGCAGGAAATGGGATCCATATAAACA	2703
Db	119801	TGTGTGCTCTATTGT-----AGGAACAGAGAGATAAAGGAGCTC	119762	Db	119123	CATACAGAACCGTCAAGATGGCAGGCTCGACTTTTGGCAGGAATGGGCCCATACAAAA	119064
Qy	1628	ATCCTCTTAGCAGATAAAGCTCTCAAACTGGGTGGACTTACTTAAAGTAAATAAAGAA	1687	Qy	2704	GGGATTTGGAATTTGCAACCAACACAGATGGGAGAAATTTCTGTGCTTACAGGTAA	2763
Db	119761	TGTAGTTACGTTAGTAAATTTCTCAAAAGTTAGTCTGATTTTC---AGTAAAGAACG	119705	Db	119063	AGGATTCGAAATATTGCCACCAATGAAGACGCGAAGAAAGTACTGTGCTTGCAGGTAA	119004
Qy	1688	TCTAATATTTGTTATATTTTCAAGGCTTATATAACACACTCCTTTAGTAACTTAG	1747	Qy	2764	CGAACAGCATGCAAAATTAATCATTTCTATTTGAAATGGGATTTTTTTTAAATTAATAA	2823
Db	119704	TCCAACTCTAGAACATATGTTGATGTTGATGATAAATGTTCTTTTACCCAGTGG	119645	Db	119003	AGACTGCGAGTGAAAAAATCAATCCCTCTATTTTAAATGAGTTTTTTTCTCAGTCAAAA	118944
Qy	1748	TAACTGTTATTTAAAGAAATGGTGAATAATCAAAAGTAAATTTCTCATAAACCCCTGAA	1807	Qy	2824	CATTCTATTGTTGGAAGCCTGTTTTAGG-----	2850
				Db	118943	CATCAAGGATGGGATTTCTTTTTTAGAATTTTAGATGCCAGGAAGATTTTAACTACAC	118884

Qy	2851	-----CAGTTAAGAGGAGTTTCCTGCACAAAAATGTGGAAGCT	2887
Db	118883	ACATGCACGTACACACATATGCACACACTAAATAAATGTGTTAAATGTGCAGCT	118824
Qy	2888	AAAGATAAGGGAAGAAAGCGAGTTTATGTGTTCCCAAATTTATTTTGGTGGAGATT	2947
Db	118823	AAGGATCTGGGAGGAA--GCTGTGCCGGTGTGCCCTCGCTTTACTTTTGTGTATGGCTT	118766
Qy	2948	TTATTTTGTGTTTCTTTTAGSTGAATATTTGGCTTGGAAATGATAAAATTAGCCAGTTAC	3007
Db	118765	ATA--TCCATTCTCTGTAGTGTGANAATTGGCTTGGCAATGATAAGATTAGCCAGCTCAC	118709
Qy	3008	CAGATGGGACCCACAGAACTTTTGATAGAAATGGAGGACTGGAAAGGAGACAAAGTAAA	3067
Db	118708	CAGGATGGGACCCACAGAACTTCTCATTTGAATGGAGGACTGGAAAGGAGACAAGGTGAA	118649
Qy	3068	GGCTCACTATGAGGATTCATGTTACAGAACTGAACCAACAAATACCAGATCTCAGTGAA	3127
Db	118648	GGCATTATGGAGGCTTCACGGTACAGACGAGGCCAGAAATACCAGATCTCAGTGAA	118589
Qy	3128	CAAAATACAGAGGAACAGCCGGTAATGCCCTCATGGATGGAGCATCTCAGCTGATGGAGA	3187
Db	118588	CAAAATACAAAGGGACGGCTGGCAACGCCCTCATGGACGGGGCTCTCAACTGGTGGGGA	118529
Qy	3188	AAA CAGGACCATGACCATTCACAAAGCGCATGTTCTTCAGCACGTAATGACAGAGACAATGA	3247
Db	118528	GAA CAGAAACCATGACCATCCACAAAGCGCATGTTCTTCAGCACGTAATGACAGAGACAATGA	118469
Qy	3248	CGGCTGGTATGTGTGGCACTCTTTGCTCTCG--CTTTAAAAATCACACTAAATCATTTACT	3306
Db	118468	CGGCTGGTAAACAACCAAGGCTCTAGCTCTGCTCTTGAATGACCCACGAAATGCGATTGTC	118409
Qy	3307	CAGAAATCAATTA--CAATATTTTAAATAGCTACCACTTCCTGGGCACTTA-----	3354
Db	118408	TAGCACTGTGGGGGTACCGAAATGTTCTAAACAACTGCCAGACTTTGGACACTTCAAGTGT	118349
Qy	3355	--CTGTGAGCACTGTGCTTAAGCTCTTTATGCACTCACTCGAAAGCATTTCAACTATAGG	3412
Db	118348	ATCAGGCAAGCCCCATCCGAACCTCCTACGTTATCAACGGGAAGCATCCAGCTATAAGA	118289
Qy	3413	TAGCAATTTCTATTCTCTCAATTTTACAGATGAGATTTTAGAGAGATTAGTGATTTGCCAAT	3472
Db	118288	CAGACATGGTTAGTTTTCATCTTTGAGATGAAGTTTAGAGATACGATTTTACCC-----	118236
Qy	3473	GTCACAACTACCCAGAGATAAAACTAGAAATTTGAGCAAGTTACTTTCTGAAATAAGA	3532
Db	118235	-TAAGACAGCTACCCAAAGATAGACCTGTATCTATAGCAACAC-----TCTGAGTTTGT	118182
Qy	3533	GCATTTAGATAAATACCTATATCTCTATATCTTAAAGTGTGTGTAAGAACTTTCAATTTTC	3592
Db	118181	GTCTCTATATACACATCTGTGTTTTCATTTCATAACGATCTACGGAAGCCATCATCCA	118122
Qy	3593	ATTTCCAGGGTTCTCTGATACTAAGGGTTGTAAGAGCTATTATTCCAGTATATAAGTAACA	3652
Db	118121	TATCTCAGCGTCT-----CTCATACTACGATATCAGTAAAG 118085	
Qy	3653	AACACAGTCCCTAGATGGAATTG-CCACAAAGCCCGAGTTATCTCTTTCTTTGCTATAGG	3711
Db	118084	AAATTCAGTCCCTTCATGATCAGATCAGATAGACCCAGCTGTTTCCGTTTCATGCTATGGG	118025
Qy	3712	GCACAGGAGGCTCTTTGGTGTATTAGTGTGACTCTATGTATAGCACCCAAAGGAAGACTTA	3771
Db	118024	GCAGAGAACAATTGGCTGTATTG---GGCTTTTACATACAGATTCAGGGAAGAGAGT	117968
Qy	3772	CTGTGCACAGAGTGTAGCAGTCTTTTATGGGTAATCTGCAAAAGCTAACTTGCACCGC	3831
Db	117967	CCACACACAGTGGAGTACAACTACTTTCATGCGGTACTCTCCAAAGGACTGAATTGATGG	117908
Qy	3832	TAGTTCTGTTTCTAATAAGCCCAACA--CATTTCTTTTCAGGTTAAATCATAGATCCCG	3889
Db	117907	TGGCACTTGATATGGATCAAACTCACTCATCTCTCTCTTTTCAGGGAATCACTGATCCAG	117848

[illegible]

Db 1817 TTATTTATGTAGGATCTGTCTCAAGAAAACTTCCAAAAAGATTATTATTAATTAACGAGACT 1876
Qy 4529 CTGTTGC 4535
| | | | |
Db 1877 CTGTTGC 1883

Search completed: November 23, 2004, 04:31:25
Job time : 20238.7 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C 1	260	12.9	1056	4	BM470567	BM470567 AGENCOURT
C 2	256.8	12.8	497	1	AL589495	AL589495 DKEZp451G
C 3	253.8	12.6	834	6	CD357520	CD357520 AGENCOURT
C 4	253.2	12.6	565	6	CA426144	CA426144 UI-H-DF0
C 5	253	12.6	576	8	AQ424282	CITBI-E1-AQ424282
C 6	252	12.5	704	5	BX645059	BX645059 DKEZp781H
C 7	250.2	12.4	635	8	B31104	CIT-HSP-217B31104
C 8	248.4	12.4	706	5	B0618120	B0618120 UI-H-DF0-
C 9	247.6	12.3	680	7	CF887062	CF887062 UI-CF-ECO
C 10	247	12.3	699	5	B0677785	B0677785 UI-CF-ECO
C 11	247	12.3	753	7	CF890780	CF890780 UI-CF-ECO
C 12	246.8	12.3	775	6	CB989598	CB989598 AGENCOURT
C 13	246.6	12.3	838	3	BC039515	Homo sapi-BC039515
C 14	246.6	12.3	1060	4	BM469250	BM469250 AGENCOURT
C 15	246.4	12.3	733	6	CB963738	CB963738 AGENCOURT
C 16	245.2	12.2	895	8	AQ514325	AQ514325 HS 5194_B
C 17	245.2	12.2	842	5	BQ437946	BQ437946 AGENCOURT
C 18	245	12.2	602	2	AW965008	AW965008 EST377081
C 19	244.8	12.2	635	1	AL691639	AL691639 DKEZp313E
C 20	242	12.0	713	9	AG011265	AG011265 Homo sapi
C 21	241.6	12.0	673	9	AG175165	AG175165 Pan trogl
C 22	241.6	12.0	3143	3	HSM805999	HSX537892 Homo sapi
C 23	241.4	12.0	642	5	BM998857	BM998857 UI-H-D10-
C 24	241.2	12.0	537	2	BF876683	QVO-ET014-BF876683

Db 526 GGTAGGCTAGTTGTGGTATTATACACCTGTAATCCACAGCACTTTGGGAGGCGCTAGGTGGG 467

Qy 348 TAGATCACTTGGAGCGCGGAGTTTGAGACACAGCTTGGCCACATGTTGTAACCCCATCTG 407

Db 466 TGGATCGTTTGGAGTCAAGGAGTTTGGAGACCAAGCTTGGCCAAACATGGTGAACCCCTGCTC 407

Qy 408 TACTAAAAATACAAAAGTTAGTCTGGCGTGTGGTGTAGATGCTCTAGTCCAGCACTACTTG 467

Db 406 TACCAAAAATATAAAAATTTAGCCAGGTGAGTGCATGTGCGCCGTAGTCCAGCACTACTTG 347

Qy 468 GGAGGCTGAGGCATGAGAATCGCTTTGAGCCGAGCTGGGCAATACAGCAAGACCCCGTCT 527

Db 346 GGAGGCTGAGGTGGGAGATCACTTGTATCCAGAGCTGCTCAACGCTGGCGAAACCCACCT 287

Qy 528 CTACAAATATAAATAAATTTAGTGTAGTGTGGTGTGATGCTGCTAGTCTCTAGCTG 587

Db 286 CTAC-----CAAAAATACAAAAATTTAGCTGGCATGTTGGTGTGACATCTGTATACAGCTA 231

Qy 588 CTAGGAGGCTGAGATGGAAGGATTTGTTGAGCTTGGGAGGTCAAGGCTGCAGTGAGCGG 647

Db 230 CTTAGGAGGCTGAGCATGAGATCACTTTGAACCCAGGAGACAGAGGTCAAGTGAGCGG 171

Qy 648 AGATGGGCCACTGCATCCAGCTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAAA 707

Db 170 ATATTGTGCCACTGCATCCAGCTGGGCAACAGAGTGAGACTCTGTCTCAAAAAAATTT 111

Qy 708 AAAAAAAGGAGGAGGAGAG 733

Db 110 AATAAATAATTAATTAATGAAAAAG 85

RESULT 2

AL589495/c

LOCUS DKEP451G0619 r1 451 (synonym: hicc1) spinal cord Homo sapiens cDNA

DEFINITION clone DKEP451G0619 5', mRNA sequence.

ACCESSION AL589495

VERSION AL589495.1 GI:13243267

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 497)

TITLE Bloecker,H., Boeher,M., Brandt,P., Mewes,W., Weil,B. and Wiemann,S.

JOURNAL EST (Bloecker,H., Boeher,M., Brandt,P., Mewes,H.W., Weil,B. and Wiemann,S.)

COMMENT Unpublished (1999)

CONTACT: MIPS

MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp451G0619) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clones@rzpd.de.

FEATURES

source

1..497

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp451G0619"

/tissue_type="human spinal cord"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="451 (synonym: hicc1) spinal cord"

/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 12.8%; Score 256.8; DB 1; Length 497;

Best Local Similarity 78.3%; Pred. No. 8.5e-32;

Matches 351; Conservative 0; Mismatches 82; Indels 15; Gaps 3;

Qy 292 GGCGGGGTGTGGTGCATGCTCTAATCCAGCACTTTGGGAGGCAAGGTGAGTAGA 351

Db 487 GGCGGGCGGGTGGTGCATGCTTACAATCCAGCACTTTGGGAGGCGGAGTGGGTGA 428

Qy 352 TCATTGAGGCGGGAGTTTGAGACCAAGCTTGGCCAAATGTTGAAACCCCATCTGTACT 411

Db 427 TCACCTGAGGTCAAGAGCTCAAGACCAAGCTTGAATCAATGGCCAAACCCCGTCTCTACT 368

Qy 412 AAAAATACAAAGTTA-GCTGGGCTGTGTGTAGATGCTGTAGTCCAGCACTTTGGGA 470

Db 367 AAAAATACAAATTTAGCCGGGCGGTGGCTCAAGCTGTATCCAGCACTTTGGGA 308

Qy 471 GGCTGAGGCATGAAATC-----GCTTGAGCCAGCTTGGCAATACAGCAAG 518

Db 307 GGCGGAGCGGGCGGATCACAAAGGTCAAGAGATCGAGACCATCTCTGGCTAACCGGTGAA 248

Qy 519 ACCCGCTCTTACAAATAAATAAATAAATAAATTTAGTGTGGTGTGATGCTGTAG 578

Db 247 ACCCGCTCTCTACTA--AAAATATAAAAAATTTAGTGTGGCGGTGGCGGCGCTGTAG 190

Qy 579 TCCTAGCTGTAGGAGGCTGAGATGGAAGGATTGCTTGAGCTGGGAGGTCAAGGCTGC 638

Db 189 TCCAGCTGTCTCCGAGGCTGAGGAGGAGATGGCATGAACCTGGGAGCGGAGCTTAC 130

Qy 639 AGTGAGCGAGATGGCGCACTGCACCTCCAGCTGGGCAACAGAGTGAGACCTGTCTCA 698

Db 129 AGTGAGCTGAGATCGCGCACTGCACCTCCAGCTGGGCGCAGAGCAAGACTCCGTTTCA 70

Qy 699 GAAAAAAGGAGGAGGAG 726

Db 69 AAAAAAAGGAGGAGGAG 42

RESULT 3

CD357520

LOCUS AGNCOURT 14248776 NIH MGC 187 Homo sapiens cDNA clone

DEFINITION IMAGE:30401007 5', mRNA sequence.

ACCESSION CD357520

VERSION CD357520.1 GI:31128931

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

834 bp mRNA linear EST 29-MAY-2003

CD357520

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 834)

NIH-MGC http://mgs.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgpbs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDCM181 row: a column: 16

High quality sequence stop: 381.

Location/Qualifiers

1..834

/organism="Homo sapiens"

/mol_type="mRNA"

Source

Query Match		12.6%;	Score 253.2;	DB 6;	Length 565;
Best Local Similarity		78.8%;	Pred. No. 3.1e-31;		
Matches 345;	Conservative	0;	Mismatches 78;	Indels 15;	Gaps 3;
<hr/>					
QY	291	TGGCCGGGTGTCGTCTCATCGCTCTAATCCAGCACATTTCGGAGGCCCAAGTGAGTAG	350		
Db	436	TGGCCGGGCATGGTGGCTTTATCGCTGTAACTCTGGCACATTTCGGAGGCCCAAGTGGGCAG	377		
QY	351	ATCACTTGAGCCGGGAGTTTCAGACCCAGCCCTGGCCCACAATCGTGTAACCCCATCTCGTAC	410		
Db	376	ATCACCCTGAGGTCAGAGTTTCAGACCCAGCCCTGGCCCCAACATGGTGAACCTCATCTCTAC	317		
QY	411	TAAAATAACAAAAGTTA-GCTCGGGGCTGGTGTAGATGCTCTAGTCCCGAGCTACTTTGGG	469		
Db	316	TAAAATAACAAAATTAGCCCGGGCGGTGGCTCAGGCTGTAACTCCAGCACATTTCGGG	257		
QY	470	AGGCTGAGGCATGAGAATC-----GCTTGAGCCCCAGCCCTGGGCAATACAGCAA	517		
Db	256	AGGCAGAGCGGGCGGATCACGAGGTCAGGAGATCAAGACCATCTCTGGCTAAACACGGTGA	197		
QY	518	GACCCCGCTCTACAAATAAATAACAAAAATTAGTTGGATGTGGTGGTGCACTGCTGTA	577		
Db	196	AACCCCGTCTCTACTA--AAATAAATAAATAATTAGCCGGGCGCTAGTGGCGGGTGGCTGTA	139		
QY	578	GTCTTAGCTGTGCTAGGAGGCTCAGATGGAAGGATTGCTTTGAGCCCTGGGAGGTCAGGGCTG	637		
Db	138	GTCCCACTACTTCGGGAGGCTGAGCAGAGAGNATGGCATGMAACCCCGGAGGCAGAGCTTG	79		
QY	638	CAGTGAGCCGAGATGGCGCCACTGCATCTCCAGCCCTGGGCAACAGAGTGAAGCCCTGTCTC	697		

```

Db      78  CAGCGAGCCGAGATTGCACCACTGCATCCAGCCTGGGCGGACAGAGCGAGACTCCGTCTC 19
Qy      698  AGAAAAAAGAAAAA 715
Db      18  AAAAAAAGAAAAA 1

RESULT 5
LOCUS   AQ424282
DEFINITION  C17BI-E1-2574H5.TR C17BI-E1 Homo sapiens genomic clone 2574H5,
genomic survey sequence.
ACCESSION  AQ424282
VERSION    AQ424282.1 GI:4497548
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 576)
AUTHORS    Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuoya,H., Simon,M. and
Venter,J.C.
TITLE      Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL   Unpublished (1997)
COMMENT   Other GSSs: C17BI-E1-2574H5.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbs@tigr.org
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES             source
    source
        1..576
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="2574H5"
            /sex="male"
            /cell_type="sperm"
            /clone_lib="C17BI-E1"
            /note="Vector: pBelBAC11; Site_1: EcoRI; Site_2: EcoRI;
            CalTech Human BAC Library D"

ORIGIN
Query Match      12.6%; Score 253; DB 8; Length 576;
Best Local Similarity 77.0%; Pred. No. 3.4e-31;
Matches 352; Conservative 0; Mismatches 90; Indels 15; Gaps 3;

Qy      287  AGCGTGGCGGGGTGTGGTGTCTTAATCCAGCACTTTGGAGGCCAAGGTGA 346
Db      38  ACGTGGCGGGCACAGTGGCTCAACCTGTAATCCAGCACTTTGGAGGCCAAGGTGG 97

Qy      347  GTAGATCACTTGAGCCGGGAGTTTGAGACACAGCTGGGCAACATGGTGAAACCCCATCT 406
Db      98  GCGGATCACTCGAGTCAAGGATTTGAGACACAGCTGGGCAAAATGGGAAACCCCTGTCT 157

Qy      407  GTACTAAAAATACAAAG-TTAGCTGGGCGTGGTGTAGTCCCTGTAGTCCAGTACT 465
Db      158  CAACATAAATGTAAAGATTAGCGGGCGTGGTGGCTCACGCTGTATCCAGCACTT 217

Qy      466  TGGAGGCTGAGGCATGGAATC-----GTTGAGCCAGCCTGGGCAATACA 513
Db      218  TGGAGGCGGAGCGGGTGAATCACAAGGTCAAGGATCGAGACCATCTCTGTAAACAG 277

Qy      514  GCAAGACCCGCTCTTACAAATAAATACAAAATATTAGTTGGATGTGGTGTGATGCC 573

```

```

Db      278  GTGAAACCTGCTCTCTACTA--AAAAATACAAAAAATTAGCCGGGAGTGGTTGTGGCGCC 335
Qy      574  TGTAGTCTTAGCTAGGAGGCTGAGATGGAAGGATTGCTTGAAGCTTGGAGGTCAAG 633
Db      336  TGTAGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCGTGAACCCGGGAGGTGGAG 395

Qy      634  GCTCAGTGAGCCGAGATGGGCGGCTGCACTTCAGCTGGGCAACAGAGTGAGACCTTG 693
Db      396  CTTGAGTGAGCGGAGATAGCACCACTGCATCTCAGGCTGGGCGACAGAGTGAGACTTTG 455

Qy      694  TCTCAGAAAAAAGAAAAAAGAAAAAAGAGAGAGAGA 730
Db      456  TCTCAAAAAAAGAAAAATTAATAAAAAATGTAAAGA 492

RESULT 6
LOCUS   BX645059/c
DEFINITION  DXFP781H2437_r1 781 (synonym: hlccc4) Homo sapiens cDNA clone
BX645059
ACCESSION  BX645059
VERSION    BX645059.1 GI:34479392
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 704)
AUTHORS    Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
TITLE      EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)
JOURNAL    Unpublished (2003)
COMMENT    Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFP781H2437) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             source
    source
        1..704
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DXFP781H2437"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="781 (synonym: hlccc4)"
            /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
            cDNA-collection"

ORIGIN
Query Match      12.5%; Score 252; DB 5; Length 704;
Best Local Similarity 78.4%; Pred. No. 4.6e-31;
Matches 345; Conservative 0; Mismatches 80; Indels 15; Gaps 3;

Qy      292  GGCGGGGTGGTGGCTCATCTTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGA 351
Db      440  GCCAGGCATGGTGGCTCACCTGTATCCAGCACTTTGGAGGCCGAGGTGAGTAGA 381

Qy      352  TCATTTAGGCGGGAGTTTGAGACCAAGCTGGGCCAAATGGTGAAACCCCATCTGTACT 411
Db      380  TCACCTTGGTCAAGGATTCGAGACCAAGCTGGGCCAAATGGTGAAACCCCATCTGTACT 321

Qy      412  AAAATA-CAAAAGTTAGCTGGCGGTGGTGTAGATGCTGTAGTCCCGAGCTACTTGGGA 470

```

```

Db      320 ATAAATAGAAAAAATGGCTGGCGTGGTGGCTCATGCTGTAATCCAGCAGCTTTGGGA 261
QY      471 GCGTGGAGCATGAGATC-----GCTTGAGCCAGCGCTGGGCAATACAGCAAG 518
Db      260 GGCAGGAGGGGGCGGATCAGAGGTGAGGATCGAGAGCATCTGGCTTAACACGGTGAA 201
QY      519 ACCCGCTCTCTACAAATAAAATACAAAAAATAGTTGGATGGTGGTGCATGCCCTGTAG 578
Db      200 ACCCGCTCTCTACTAAAAATACAAAAAATAGCCGGCTGGTGGCGGCGCTGTAG 141
QY      579 TCTAGCTGCTAGGAGGCTGAGATCGAAGGAT--TGCTTGAGCCCTGGGAGGTCAAGGCT 636
Db      140 TCCAGCTACTCGGAGGCTGAGGACGAGAGATGGTGGTGAAACCCGGGAGGGAGCTT 81
QY      637 GCAGTGAGCGGAGATGGCGCCACTGCCTCCAGCTGGGCAACAGAGTGAGACCCCTGTCT 596
Db      80 GCAGTGAGCGGAGATGCGACCACTGCCTCCAGCTGGGCGGACAGAGCGAGATCCATCT 21
QY      697 CAGAAAAAATAAAAAA 716
Db      20 CAAAAAATAAAAAA 1

```

```

RESULT 7
B91104/c
LOCUS   B91104
DEFINITION CIT-HSP-2173120.TF CIT-HSP Homo sapiens genomic clone 2173120,
Genomic survey sequence.
ACCESSION B91104
VERSION   B91104.1 GI:2973584
KEYWORDS  GSS.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

```

REFERENCE 1 (bases 1 to 635)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING  Building (1998)
JOURNAL   Unpublished (1998)
COMMENT   Other GSSs: CIT-HSP-2173120.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.

```

```

FEATURES             Location/Qualifiers
     source            1..635
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="GDB:7104270"
                        /db_xref="taxon:9606"
                        /clone="2173120"
                        /sex="Male"
                        /cell_type="Sperm"
                        /clone_lib="CIT-HSP"
                        /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
                        HindIII"
ORIGIN
Query Match      12.4%; Score 250.2; DB 8; Length 635;
Best Local Similarity 73.5%; Pred. No. 9.3e-31;
Matches 352; Conservative 0; Mismatches 113; Indels 14; Gaps 2;

```

```

QY      245 CTACAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCCGGTGTGGT 304
Db      635 CTGTAGCATAGAGAAACATCCACCTTTTAGTGCAAAAAGTTGTGGCCAAATGCAGT 576
QY      305 GGCTCATGCTCTAATCCAGCACATTTGGAGGCCCAAGGTGAGTAGATCACTTTGAGGCCG 364
Db      575 GGCTCAAGCTGTATCCAGCACATTTGGGAAGCCGAGGTGGACCGATCACTTGAGGTCA 516
QY      365 GGAGTTTGAGACAGCGCTGGCCAAACATGGTGAAGACCCCATCTGTACTAAAAATACAAAAG 424
Db      515 GGAGTTTGAAAACAGCGCTGGCCAAACATGGCAAAACCCCATCTCTACTAAAAATACAAAAG 456
QY      425 TTAGCTGGCGGTGGTGTAGATGCTGTAGTCCAGCTACTTTGGGAGGCTGGAGGATGAG 484
Db      455 TTAGCCAGGCACGGTGGTGGCTTCCCAATCCAGCACTTTCCGAGGCTGAGGGGGCA 396
QY      485 AATCGCTTG-----AGCCAGCGCTGGGCAATACAGCAAGACCCCGTCTCTACA 532
Db      395 GATCACAGGTGAGGAGGACCAAGACCATCTCGCTAACACAGTGAACCCCATCTCTACT 336
QY      533 AATAAAATACAAAAATTAGTTGGATGGTGGTGCATGCCCTGTAGTCTTAGCTGTAGG 592
Db      335 A--AAATACAAAAAATTAGCCGGGTGTGGTGGCGGCACCTGTAGTCCACTCTCAG 278
QY      593 GAGGCTGAGATCGAAGGATGCTTGAGCTGGAGGTCAAGCTGAGCTGAGTGGCCGAGATG 652
Db      277 GAGGCTGAGGCGAGGAATGGATGAACCCGGGAGCGGAGCTTGCAAGTGGAGATT 218
QY      653 GCGCCACTGCACTCCAGCGCTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAA 711
Db      217 GCGCCACTGCACTCCAGCGCTGGGCAACAGTGTGAGACTCCGTCTCAAAAAA 159

```

```

RESULT 8
B918120/c
LOCUS   B918120
DEFINITION UI-H-DF0-beu-1-08-0-UI.s1 NCI_CGAP_DFO Homo sapiens cDNA clone
UI-H-DF0-beu-1-08-0-UI 3', mRNA sequence.
ACCESSION B918120
VERSION   B918120.1 GI:23284335
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 11-299, >ALU (matched complement) 217-435, >ALU (matched
complement) 351-572, >ALU (matched complement) 574-633,
>MIR#SINE/MIR
Seq primer: M13 FORWARD
POLYA=Yes.

```

```

FEATURES             Location/Qualifiers
     source            1..706
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="UI-H-DF0-beu-1-08-0-UI"
                        /tissue_type="Subchondral Bone"
                        /dev_stage="Adult"

```



```

Db      507 CTACTGAGAGGCTGAGCGCAGGAGATGCGCGCAACTGGGAGCGGAGCTTGCACTGAG 566
QY      645 CCGAGATGGCGCCACTGCACTCCAGCCTGGGCAACAGAGTGAGACCCCTGCTCAGAA--- 701
Db      567 CCAAGATCGGCACCTGCACTCCAGCCTGGGTCAGAGAGTGAGACTCCATCTCNAAAAC 626
QY      702 AAAAAAAAAAAAAAAAAAGGAGGAGAGAGACTCAAGCACCCCTCCACAGGACTGCT 761
Db      627 AAAAAACAAAAACAAAAACAGCTAAGACTTTCAACTAAGCATCANGATGGGAGGAG 686
QY      762 GAGGCCCTGAGGTGTCTGAGCATGTGGCCCGAGCGGGGAGACTCTGTAAAGCC 815
Db      687 GAGGTGTG-GGGTTTTTGAAGGAAGTGACCTCAAGCTGAGCACCCCTGTTTCCC 739

RESULT 13
BC039515
LOCUS   BC039515      838 bp      mRNA      linear      HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:5585187, mRNA.
ACCESSION BC039515
VERSION   BC039515.1 GI:24659995
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Strausberg, R.
          Direct Submission
          Submitted (01-NOV-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
          NIH-MGC Project URL: http://mgc.nci.nih.gov
          Contact: MGC help desk
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Sequencing Group at the Stanford Human Genome
          Center, Stanford University School of Medicine, Stanford, CA 94305
          Web site: http://www.shgc.stanford.edu
          Contact: (Dickson, Mark) mcd@paxil.stanford.edu
          Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
          R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Length: 88 Row: h Column: 14
This clone has the following problem: retained intron.

FEATURES             source
          Location/Qualifiers
          1..838
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5585187"
             /tissue_type="Testis, embryonal carcinoma"
             /clone_lib="NIH_MGC_92"
             /lab_host="DH10B"
             /note="Vector: pCMV-SPORT6"

ORIGIN
Query Match      12.3%; Score 246.6; DB 3; Length 838;
Best Local Similarity 74.1%; Pred. No. 3.3e-30;
Matches 358; Conservative 0; Mismatches 109; Indels 16; Gaps 3;

QY      239 TTGGCTCAAGATTGTGGATGGGTACCCAGTGCACAGTGAGGGAGGCTGCGCGG 298
Db      168 TGTGGTCTGAAATCCTCAGCTGATACTTAGTTTAAATAGTCTCAAGTAGGCTGG 227
QY      299 TGTGGTGGCTAGCTCTTAATCCAGCACTTTGGAGGCCCAAGGTGATAGATCACTTG 358
Db      228 CGTGGTGGCTACACCTGTGAATCCCACTTTGGAGGCCGAGCGGATCACCTG 287

```

```

QY      359 AGCGCGGAGTTTGACACAGCCTGGCCAAACATGGTGAACCCCATCTGTACTAAAAATA 418
Db      288 AGGTACAGGAGTTTCAAGACAGCCTGGCCAAACATGGTGAACCCCATCTGTACTAAAAATA 347
QY      419 CAAAGTTAGCTGGGGCTGGTGTAGATCCCTGTAGTCCAGCTACTTTGGGAGGCTGAGG 478
Db      348 CAAAA--TTAGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 406
QY      479 CATGAGAATC-----GCTTGAGCCCGAGCTGGGCAATACAGCAAGACCCCGTC 526
Db      407 CAGGTGATCAGAGGCTCAGGGGTTTCGAGACCAAGCTGACCAACATGGTGAACCCCTGTC 466
QY      527 TCTCAAAATAAAATACAAAAAATTTAGTTGGATGTGGTGGTGGTGGTGGTGGTGGTGGTGG 586
Db      467 TCTAC---TAAAAAATACAAAAAATTTAGCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 523
QY      587 GCTAGGAGGCTGAGATGAAGATTGCTTTGAGCCTGGGAGGTTCAAGGCTGCAAGTCAGGC 646
Db      524 ACTCGGAGGCTGAAGCGGAGAAATTGCTTGAACGGGAGGGGAGGTTGCAAGTGAACC 593
QY      647 GAGATGGCGCCACTGCACTCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAAA 706
Db      584 AGGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCAGACTCGTCTCAAAAAAAA 643
QY      707 AAA 709
Db      644 AAA 646

BM469250      1060 bp      mRNA      linear      EST 05-FEB-2002
AGENCOURT_6444645 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5585187
5', mRNA sequence.
BM469250
BM469250.1 GI:18518292
EST.
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          NIH-MGC http://mgc.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM12351 row: b column: 04
          High quality sequence stop: 659.
          Location/Qualifiers
          1..1060
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5585187"
             /tissue_type="embryonal carcinoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_92"
             /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
             Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
             Average insert size 2.5 kb. Library enriched for
             full-length clones and constructed by Life Technologies.
             Note: this is a NIH_MGC Library."

FEATURES             source
          Location/Qualifiers
          1..1060
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5585187"
             /tissue_type="embryonal carcinoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_92"
             /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
             Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
             Average insert size 2.5 kb. Library enriched for
             full-length clones and constructed by Life Technologies.
             Note: this is a NIH_MGC Library."

ORIGIN

```

Query Match 12.3%; Score 246.4; DB 4; Length 1060;
 Best Local Similarity 74.1%; Pred. No. 3.1e-30;
 Matches 358; Conservative 0; Mismatches 109; Indels 16; Gaps 3;

QY 239 TTTGGCTACAAAGTTGTGGATGTGGTRACCATGCCAAGTGTGAGGGAGAGCTGGCCGGG 298
 DB 168 TGTGCTGTGAAAAATCCTCAGGCTGATAACTTAGTTTAAAAATAGTGTCTCAAGTAGGCTGGG 227

QY 299 TGTGCTGCTCATGCTCTTAATCCAGCACATTTGGAGGCCAAGGTAGTAGTACATCCTG 358
 DB 228 CGTGTGCTGTACACCTGTATCCACACATTTGGAGGCCGAGGCGGATCACCTG 287

QY 359 AGCGCGGAGTTTGAGACACGCTGGCCAAACATGTGAAACCCCATCTGTACTAAAAATA 418
 DB 288 AGGTCAAGAGTTCAAGACACGCTGGCCAAACATGTGAAACCCCATCTGTACTAAAAATA 347

QY 419 CAAAAGTTAGCTGGCGTGTGTAGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGG 478
 DB 348 CAAAATTTAGCTGGGTGTGTGTGGTGTGCTGTAAATCCAGCATTTGGAGGCGAGG 406

QY 479 CATGAGATC-----GCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTC 526
 DB 407 CAGGTGATCAGAGGTCAAGGCTGAGAGCCAGCTGACCAACATGTGAAACCTGTC 466

QY 527 TCTACAAATAAAATAACAAAATTTAGTTGGATGTGGTGTGATGCCCTGTAGTCTAGCT 586
 DB 467 TCTAC---TAAATAACAAAATTTAGCCGGTGTGGTGGGCGCTGTAAATCCAGCT 523

QY 587 GCTAGGAGGCTGAGATGGAAGATGCTTGAGCTTGGAGGTCAAGCTGCAAGTGAGCC 646
 DB 524 ACTCGGAGGCTGAAGCGGAGAAATTTGCTTGAACCGGAGCGGAGGTGTCAGTGAACC 583

QY 647 GAGATGGCGCCACTCCTCCAGCTGGCAACAGAGTGAGACCTGCTCTCAGAAAAAAA 706
 DB 584 AGGATCGCGCCACTGCTCCAGCTGGCGAGAGGAGACTCGCTCTCAAAAAAAA 643

QY 707 AAA 709
 DB 644 AAA 646

RESULT 15
 CB963738
 LOCUS
 DEFINITION AGENCOURT 13464887 NIH MGC 187 Homo sapiens cdna clone
 IMAGE:30320300 5', mRNA sequence.
 ACCESSION CB963738
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 733)
 NIH-MGC http://mgc.mci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDCM117 row: n column: 21
 High quality sequence stop: 550.

FEATURES
 source
 1..733
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

/clones="IMAGE:30320300"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lib="NIH MGC 187"
 /note="Organ: Blood vessels - aorta, basilar and artery;
 Vector: pBNR-LIB; Site 1: SfiI (ggcattatggcc); Site 2:
 SfiI (ggcgctcgcc); 5' and 3' adaptors were used in
 cloning as follows: 5' adaptor sequence:
 5'-CACGCGCATATGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGGCGGCGCATG-DT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.4 kb
 (range 0.5-4.0 kb). 14/15 colonies contained inserts by
 PCR. This library was enriched for full-length clones and
 was constructed by Clontech Laboratories (Palo Alto, CA).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 12.3%; Score 246.4; DB 6; Length 733;
 Best Local Similarity 76.9%; Pred. No. 3.7e-30;
 Matches 332; Conservative 0; Mismatches 86; Indels 14; Gaps 2;

QY 294 CCGGCTGTGGTGTCTATCTTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATC 353
 DB 99 CTGGCGCAGTGTTCACGCTGTAAATCCAGCACTTTGGAGGCCAAGGGGTGATC 158

QY 354 ACTTGAGCCCGGAGTTTGACACCAAGCTGGCCACATGTTGAACCCCATCTGTACTAA 413
 DB 159 ACCTGAGTTGGGAGTTTCGAGACCAGCCTGGCCAAACATGGTGAACCCCTGTCTACTGA 218

QY 414 AAATACAAAAGTTAGCTGGCGTGGTGTAGATGCTGTAGTCCAGCTACTTTGGAGGC 473
 DB 219 AAATACAAAAGTTAGCTGGTGTGGTGTGACACACCTGTAGTCCCATCTTTGGAGGC 278

QY 474 TGAGGCATGAGATC-----GCTTGAGCCAGCTGGGCAATACAGCAAGACC 521
 DB 279 TGAGCGGCTGGATCAGCAGGTTCAGAGATCGAGACCATCTTGGCTAACATGGTGAACCC 338

QY 522 CGCTCTTACAAATAAAATACAAAATTTAGTTGGATGTGGTGGTGCATGCTGTAGTCC 581
 DB 339 CCATCTCTACTA--AAAAATACAAAATTTATCTGGCGGTGGTGGCGTGGCGCTGTAGTCC 396

QY 582 TAGCTGTAGGAGGCTGAGATGGAAGGATTGCTTGAGCTGGGAGGTCAAGGCTGCAGT 641
 DB 397 CAGCTACTCAGAGGTTGAGGAGGAGATCGCTTGAACCCAGAGGAGAGGTTGCAGT 456

QY 642 GAGCCGAGATGGCGCCACTGCACTCCAGCTGGGCAACAGAGTGAGAGCCCTGTCTCAGAA 701
 DB 457 GAGCCGAGATTGGCGCCACTGCACTCCAGCTGGGTGACAGAGCAAGACTCCACCTCAAAG 516

QY 702 AAAAAA 713
 DB 517 TAATAATAATA 528

Search completed: November 23, 2004, 12:26:22
 Job time : 5943.85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:20 ; Search time 890.868 Seconds
(without alignments)
11849.778 Million cell updates/sec

Title: US-10-017-724-3_COPY_85400_87410

Perfect score: 2011

Sequence: 1 ccggccgcttgaagagctt.....gaatgtggaagtaggcggt 2011

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2011	100.0	98229	9	AAD55116 Human ACE
2	1233.6	61.3	2849	2	AAG35034 DNA fragm
3	824.6	41.0	1856	3	AAG38329 Human ang
4	824.6	41.0	1856	10	ADF39358 Intron 16
5	599.4	29.8	616	10	ADC26746 Human lip
6	270.4	13.4	41369	10	ADL13522 Osteoartr
7	267.2	13.3	2836	11	ADM01345 Human CDN
8	267	13.3	12730	4	AAK70514 Human imm
9	265.4	13.2	300000	10	ADE86352 Human PTP
10	265.4	13.2	300001	12	ADO14076 Human pro
11	262.8	13.1	26874	8	AAK67255 Human CCN
12	262.8	13.1	26874	9	ADA02467 Human CCN
13	262.8	13.1	26874	10	ADB72206 Human CCN
14	262	13.0	20951	6	AAU53468 Genomic D
15	261.4	13.0	2568	6	AAK99739 cDNA enco
16	260.8	13.0	3769	4	AAK67255 Human imm
17	259.2	12.9	1496	6	AAU50885 Human Mch
18	258.6	12.9	3276	5	AAU50885 Human Mch
19	258.6	12.9	3276	10	ADB33435 Human nov
20	257.8	12.8	4301	4	AAI68146 Human agg
21	257.8	12.8	4301	10	ADL13487 Osteoartr

C 22	257.8	12.8	4303	3	AAA95826 Human met
C 23	257.8	12.8	4307	10	ADB85504 Human egg
C 24	257.8	12.8	4307	10	ACF63399 Human ADA
C 25	257.8	12.8	4406	4	AAS46100 Human DNA
C 26	257.8	12.8	4407	3	AAA37111 Human PRO
C 27	257.8	12.8	4407	4	AAF54426 Probe #46
C 28	257.8	12.8	4407	8	ACA89550 cDNA enco
C 29	257.8	12.8	4407	8	ACA73560 Human sec
C 30	257.8	12.8	4407	8	ACA05875 Human sec
C 31	257.8	12.8	4407	8	ACA66709 cDNA enco
C 32	257.8	12.8	4407	8	ACF20284 Human sec
C 33	257.8	12.8	4407	8	ACF19670 Human sec
C 34	257.8	12.8	4407	8	ACD21958 Human sec
C 35	257.8	12.8	4407	8	ACF13123 Human sec
C 36	257.8	12.8	4407	8	ACD25226 Human sec
C 37	257.8	12.8	4407	8	ACF00275 Human sec
C 38	257.8	12.8	4407	8	ACA72332 Novel hum
C 39	257.8	12.8	4407	8	ACD04856 Novel hum
C 40	257.8	12.8	4407	8	ACD18317 Human sec
C 41	257.8	12.8	4407	8	ACD08324 Human sec
C 42	257.8	12.8	4407	8	ACA88758 Novel hum
C 43	257.8	12.8	4407	8	ACA70200 Human sec
C 44	257.8	12.8	4407	8	ACD12422 Novel hum
C 45	257.8	12.8	4407	8	ACC74337 Human sec

ALIGNMENTS

RESULT 1

AAD55116

ID AAD55116 standard; DNA; 98829 BP.

XX

AC AAD55116;

DT 07-AUG-2003 (first entry)

XX Human ACE reference DNA (GI 13027555).

XX Thrombospondin 2; THBS2; angiotensin converting enzyme; polymorphism;
ACE-1; beta-fibrinogen; FGB; peripheral vascular disease; ischaemia;
KW vascular disease; myocardial infarction; pulmonary embolism; stroke;
KW atherosclerosis; coronary artery disease; venous thromboembolism; human;
KW gene; ds.
XX Homo sapiens.
OS WO2003020118-A2.
PN 13-MAR-2003.
PD 04-SEP-2002; 2002WO-US028113.
PF 05-SEP-2001; 2001US-0317178P.
XX 16-OCT-2001; 2001US-0329958P.
PR 14-DEC-2001; 2001US-00017724.
XX (VITI-) VITIVITY INC.
PA Mccarthy J;
XX WPI; 2003-300816/29.
DR Identifying polymorphisms in thrombospondin 2, angiotensin converting
XX enzyme and/or beta-fibrinogen genes in nucleic acid sample of subject, by
PT contacting the nucleic acid with a complementary probe or primer.
XX Claim 3; Fig 3; 194pp; English.
XX The invention relates to a method for determining the identity of one or
CC more allelic variants of a polymorphic region of a thrombospondin 2
CC (THBS2), angiotensin converting enzyme (ACE)-1 and/or beta-fibrinogen
CC (FGB) genes in a nucleic acid obtained from a subject. The method

CC involves contacting the nucleic acid with a complementary probe or
CC primer. The method is useful for diagnosing or aiding in the diagnosis of
CC vascular disease or disorder in a subject e.g. myocardial infarction,
CC coronary artery disease, atherosclerosis, ischemia, stroke, peripheral
CC vascular disease, venous thromboembolism and pulmonary embolism. The
CC present sequence is human ACE reference DNA. Note: This sequence is said
CC to encode SEQ ID NO: 4 (AAE36412). However this does not appear to be the
CC case
XX
SQ Sequence 98829 BP; 21920 A; 26068 C; 26526 G; 23318 T; 0 U; 997 Other;
Query Match 100.0%; Score 2011; DB 9; Length 98829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGGCCCTTTGAAGAGACTACCCCGACACAGGACCCGACACATACCTCCGAGC 60
DB CGGGCCCTTTGAAGAGACTACCCCGACACAGGACCCGACACATACCTCCGAGC 85459
QY 61 TCCCTCTCAACCCACCCCTTCCAGGTTGGAGAACTTGAGGCATAACTTGCTTCCATGA 120
DB TCCCTCTCAACCCACCCCTTCCAGGTTGGAGAACTTGAGGCATAACTTGCTTCCATGA 85519
QY 121 GGAATCTCAACCCAGAAATGGGTCTTCTGGCCCCCAGCCAGCTCCCAATTAGAA 180
DB GGAATCTCAACCCAGAAATGGGTCTTCTGGCCCCCAGCCAGCTCCCAATTAGAA 85579
QY 181 TGCAAAATAGAGGGGAATGGAATAAAGAGGAGAACGGTTTCCAGGACAGGGTT 240
DB TGCAAAATAGAGGGGAATGGAATAAAGAGGAGAACGGTTTCCAGGACAGGGTT 85639
QY 241 TGGCCCTCAAGTTGTGGATGTGGGTACCCATGCAAGTGTAGGGGAGGCTGGCCGGGTG 300
DB TGGCCCTCAAGTTGTGGATGTGGGTACCCATGCAAGTGTAGGGGAGGCTGGCCGGGTG 85699
QY 301 TGGTGGCTATGCTCTTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTGAG 360
DB TGGTGGCTATGCTCTTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTGAG 85759
QY 361 GCGGGAGTTTGAGACGACCTGGCCCAACATGGTGAACCCCATCTGTACTAAATAATACA 420
DB GCGGGAGTTTGAGACGACCTGGCCCAACATGGTGAACCCCATCTGTACTAAATAATACA 85819
QY 421 AAAGTTAGTGGCGTGTGGTAGATGCTGTAGTCCAGCTACCTTGGGAGGCTGAGGCA 480
DB AAAGTTAGTGGCGTGTGGTAGATGCTGTAGTCCAGCTACCTTGGGAGGCTGAGGCA 85879
QY 481 TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCTCTCTCAAAATAAAT 540
DB TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCTCTCTCAAAATAAAT 85939
QY 541 ACAGAAATATAGTTGGATGTGGTGTGCATGCTGTAGTCTGTAGTCTAGGAGGCTGA 600
DB ACAGAAATATAGTTGGATGTGGTGTGCATGCTGTAGTCTGTAGTCTAGGAGGCTGA 85999
QY 601 GATGAAGGATGTGCTTGAGCTGGAGGTCAAGGCTGCAAGTGCAGCGGAGATGGCCACT 660
DB GATGAAGGATGTGCTTGAGCTGGAGGTCAAGGCTGCAAGTGCAGCGGAGATGGCCACT 86059
QY 661 GCACTCCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAATAAAAAA 720
DB GCACTCCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAATAAAAAA 86119
QY 721 GGAGAGGAGAGAGACTCAAGCAGCCCTTCAAGAGCTGCTCAGGCGCTGAGGTGTCTG 780
DB GGAGAGGAGAGAGACTCAAGCAGCCCTTCAAGAGCTGCTCAGGCGCTGAGGTGTCTG 86179
QY 781 CAGCATGTGGCCCGAGCGGGGACTCTGTAAAGCACTGCTGGAGAGCACTCCCATCT 840
DB CAGCATGTGGCCCGAGCGGGGACTCTGTAAAGCACTGCTGGAGAGCACTCCCATCT 86239
QY 841 TTCTCCCATTTCTCTAGACTGCTGCTTATACAGTCACTTTTATGTGGTTTCCCAATTT 900
DB TTCTCCCATTTCTCTAGACTGCTGCTTATACAGTCACTTTTATGTGGTTTCCCAATTT 87379

DB 86240 TTCTCCCATTTCTCTAGACTGCTGCTTATACAGTCACTTTTATGTGGTTTCCCAATTT 86299
QY 901 TATTCCAGCTCTGAAATTTCTCTGAGCTCCCTTAAAGCAGAGGTGAGCTAAGGCTCGA 960
DB TATTCCAGCTCTGAAATTTCTCTGAGCTCCCTTAAAGCAGAGGTGAGCTAAGGCTCGA 86359
QY 961 GCTCAAGGCATTTAAACCCCTTACCAGATCTGACGAATGTGATGTCGACCGTCCCGAATA 1020
DB GCTCAAGGCATTTAAACCCCTTACCAGATCTGACGAATGTGATGTCGACCGTCCCGAATA 86419
QY 1021 TGAAGACTTGTATGGGATGGAGGCTGCGGAGACAGCGGGGAGAGCACTCTTCCA 1080
DB TGAAGACTTGTATGGGATGGAGGCTGCGGAGACAGCGGGGAGAGCACTCTTCCA 86479
QY 1081 GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTCCCGGCTCAATGTGAGTCCCT 1140
DB GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTCCCGGCTCAATGTGAGTCCCT 86539
QY 1141 GCTGCCAATCATCTGCGCACTTGGGTCCCTTCAATTTTCTCAAGAGGTGCTGTGAAC 1200
DB GCTGCCAATCATCTGCGCACTTGGGTCCCTTCAATTTTCTCAAGAGGTGCTGTGAAC 86599
QY 1201 CCAAGCTTAGAAAGGTAGATCCCTGAGAGGAGGAGTATGTGCTTGGAGAGCT 1260
DB CCAAGCTTAGAAAGGTAGATCCCTGAGAGGAGGAGTATGTGCTTGGAGAGCT 86659
QY 1261 GGCTGTGCTCCCTCTGTAGGCTATGTAGATGAGGGGACTCGTGAGGCTTATGTACGAG 1320
DB GGCTGTGCTCCCTCTGTAGGCTATGTAGATGAGGGGACTCGTGAGGCTTATGTACGAG 86719
QY 1321 ACACCATCCCTGGAGCAAGACTGAGGCGCTTTTCCAGGAGCTGACGCCACTTACCTC 1380
DB ACACCATCCCTGGAGCAAGACTGAGGCGCTTTTCCAGGAGCTGACGCCACTTACCTC 86779
QY 1381 AACCTGCATGCTAGTCCGCGGCGCTGACCGTCACTACGCGGCCAGCACATCAAC 1440
DB AACCTGCATGCTAGTCCGCGGCGCTGACCGTCACTACGCGGCCAGCACATCAAC 86839
QY 1441 CTGAGGGGGCCATTTCTGCTCACCTGTGGGTAAAGGCACATGTGCGGGCTTGAAGAG 1500
DB CTGAGGGGGCCATTTCTGCTCACCTGTGGGTAAAGGCACATGTGCGGGCTTGAAGAG 86899
QY 1501 GTAAAGACGACCAAGTGTGAGGTGGGACAGGCTGACTAGAGGTGAGGAGC 1560
DB GTAAAGACGACCAAGTGTGAGGTGGGACAGGCTGACTAGAGGTGAGGAGC 86959
QY 1561 AGGCTGGGACCTGAGAGACTCCAGCCCTGTGGGGATGTTGCCAGGCTGGAGGGGCT 1620
DB AGGCTGGGACCTGAGAGACTCCAGCCCTGTGGGGATGTTGCCAGGCTGGAGGGGCT 87019
QY 1621 GGGGCTGGGAGTGGGAGGCCCCCATTGATCTGTGGTCCCACTTCACTGAGATCTAT 1680
DB GGGGCTGGGAGTGGGAGGCCCCCATTGATCTGTGGTCCCACTTCACTGAGATCTAT 87079
QY 1681 GTGCGGCAAGTCAACATGGATGGGGAAGAGTTAATCTTGTTCAGGAGACCAAGGC 1740
DB GTGCGGCAAGTCAACATGGATGGGGAAGAGTTAATCTTGTTCAGGAGACCAAGGC 87139
QY 1741 ACCCATCAACATTTGTGATCTTAGAGGGGAGGAGAGGCTGTGAGTGGGAGCTGG 1800
DB ACCCATCAACATTTGTGATCTTAGAGGGGAGGAGAGGCTGTGAGTGGGAGCTGG 87199
QY 1801 GAGGCTTTTCCAAAGAGTGGGCTGTGAGCAGGGCTCGGAGAGTACAGGGTTGACAGA 1860
DB GAGGCTTTTCCAAAGAGTGGGCTGTGAGCAGGGCTCGGAGAGTACAGGGTTGACAGA 87259
QY 1861 TGGGAAGTGGGGATGAGAGGACAGCGAGTGTTCAGGCCAGGGAACCTGAAACAGA 1920
DB TGGGAAGTGGGGATGAGAGGACAGCGAGTGTTCAGGCCAGGGAACCTGAAACAGA 87319
QY 1921 AGAACCTGAGAAATGTAATCTACTTCAACCTTGGAGCCCTCTTTCGCAAGGGCTGCAATC 1980
DB AGAACCTGAGAAATGTAATCTACTTCAACCTTGGAGCCCTCTTTCGCAAGGGCTGCAATC 87379

QY 1981 TCAGATGCCCTGAATGTGTAAGTAGCGGT 2011
 |||||
 Db 87380 TCAGATGCCCTGAATGTGTAAGTAGCGGT 87410
 |||||
 RESULT 2
 AAQ35034
 ID AAQ35034 standard; DNA; 2649 BP.
 XX
 AC AAQ35034;
 XX
 DT 25-MAR-2003 (revised)
 DT 27-MAY-1993 (first entry)
 XX
 XX DNA fragment contg. ACE gene intron 16.
 XX
 KW Angiotensin converting enzyme; polymorphism; detection; insertion;
 KW vasoactive peptide metabolism; disease; hypertension; diagnosis;
 KW atherosclerosis; granulomatous; sugar diabetes complications; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT intron 1..82
 FT /*tag= a
 FT /note= "end of intron 15"
 FT exon 83..170
 FT /*tag= b
 FT /note= "exon 16"
 FT intron 171..2026
 FT /*tag= c
 FT /note= "intron 16"
 FT misc_feature 1621..1908
 FT /*tag= f
 FT /note= "repetitive Alu-type insertion sequence"
 FT exon 2027..2170
 FT /*tag= d
 FT intron 2171..2649
 FT /*tag= e
 FT /note= "intron 17"
 XX
 PN WO9300360-A1.
 XX
 XX 07-JAN-1993.
 XX
 XX 24-JUN-1992; 92MO-FR000574.
 XX
 XX 27-JUN-1991; 91PR-00008020.
 XX
 XX (INERM) INSERM INST NAT SANTE & RECH MED.
 PA
 PI Soubrier F, Hubert C, Corvol P;
 XX
 XX WPI; 1993-036331/04.
 DR
 XX
 XX Detection of polymorphism in the angiotensin-I converting enzyme gene -
 PT using new DNA sequences, for diagnosing atherosclerosis, granuloma-
 PT associated diseases and diabetic complications.
 XX
 XX Claim 1; Fig 2; 22pp; French.
 XX
 CC The sequence is that of a DNA fragment contg. intron 16 of the gene for
 CC angiotensin-I converting enzyme (ACE). Primers based on part of the
 CC sequence of the intron can be used to detect polymorphism in the ACE
 CC gene, esp. by hybridisation or amplification (partic. PCR), and to detect
 CC the presence/absence of a defined insertion (i.e. nucleotides 1451-1738
 CC of the intron sequence). This can be used to study the impact of
 CC polymorphism on diseases, metabolism of vasoactive peptides and
 CC mechanisms of hypertension. In particular it can be used in diagnosis of
 CC atherosclerosis, granulomatous diseases and the complications of sugar
 CC diabetes. (Updated on 25-MAR-2003 to correct PN field.)

XX
 SQ Sequence 2649 BP; 580 A; 761 C; 733 G; 575 T; 0 U; 0 Other;
 Query Match 61.3%; Score 1233.6; DB 2; Length 2649;
 Best Local Similarity 84.0%; Pred. No. 7.4e-309;
 Matches 1594; Conservative 0; Mismatches 9; Indels 295; Gaps 6;
 QY 1 CCGGGCCGTTTGAAGAGCTCACCCGACACAGGACCCGACACAGATACCTCCAGC 60
 |||||
 Db 759 CCAGCCCGTTTGAAGAGCTCACCCGACACAGGACCCGACACAGATACCTCCAGC 818
 |||||
 QY 61 TCCCTCTCAACCCACCCCTTTCAGGGTTGGAGAACTTGAGGCATAAATCTTCTCCATGA 120
 |||||
 Db 819 TCCCTCTCAACCCACCCCTTTCAGGGTTGGAGAACTTGAGGCATAAATCTTCTCCATGA 878
 |||||
 QY 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCAGCCCCAGCTCCCAATPAGAACAA 180
 |||||
 Db 879 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCAGCCCCAGCTCCCAATPAGAACAA 938
 |||||
 QY 181 TGACAAATAGAGGGGAATGGAATAAACAAGGAGAAAGGTTTCCAGGACAGGGTT 240
 |||||
 Db 939 TGACAAATAGAGGGGAATGGAATAAACAAGGAGAAAGGTTTCCAGGACAGGGTT 998
 |||||
 QY 241 TGGCCCTACAAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGGAGGCTGCCCGGGTG 300
 |||||
 Db 999 TGGCCCTACAAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGGAGGCTGCCCGGGTG 1058
 |||||
 QY 301 TGTGGCTCATGCCCTTAATCCAGACATTTGGGAGGCCAAGGTGAGTAGATCACTTGAG 360
 |||||
 Db 1059 TGTGGCTCATG-CTCTAATCCAGACATTTGGGAGGCCAAGGTGAGTAGATCACTTGAG 1117
 |||||
 QY 361 GCCGGGAGTTTGAGACCCAGCTGGCCACATGTCGAACCCCATCTGTACTAATAATACA 420
 |||||
 Db 1118 GCCGGGAGTTTGAGACCCAGCTGGCCACATGTCGAACCCCATCTGTACTAATAATACA 1177
 |||||
 QY 421 AAAGTTAGCTGGGCGTGGTGTAGATGCTGTAGTCCCAAGTACTTTGGAGGCTGAGGCA 480
 |||||
 Db 1178 AAAGTTAGCTGGGCGTGGTGTAGATGCTGTAGTCCCAAGTACTTTGGAGGCTGAGGCA 1237
 |||||
 QY 481 TGAAATCGCTTGAGCCAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 540
 |||||
 Db 1238 TGAAATCGCTTGAGCCAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 1297
 |||||
 QY 541 ACAGAAATATTAGTTGGATGTGGTGTGATGCTGTAGTCCCTAGCTGTAGGAGGCTGA 600
 |||||
 Db 1298 ACAGAAATATTAGTTGGATGTGGTGTGATGCTGTAGTCCCTAGCTGTAGGAGGCTGA 1357
 |||||
 QY 601 GATGAAGGATTGCTTTGAGCCCTGGGAGGTCAAGGCTGCAGTGCAGCCGAGATGGCGCACT 660
 |||||
 Db 1358 GATGAAGGATTGCTTTGAGCCCTGGGAGGTCAAGGCTGCAGTGCAGCCGAGATGGCGCACT 1417
 |||||
 QY 661 GCATCTCCAGCTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAATAAAAAAATAAAA 720
 |||||
 Db 1418 GCATCTCCAGCTGGGCAACAGAGTGAGACCCCTGTCTCAG--AAAGAAAAAATAAAAAA 1475
 |||||
 QY 721 GGAGAGGAGAGAGACTCAAGCAGCCCTCAGCAGACTGCTCAGGCCCTGAGGTTCTG 780
 |||||
 Db 1476 GGAGAGGAGAGAGACTCAAGCAGCCCTCAGCAGACTGCTCAGGCCCTGAGGTTCTG 1535
 |||||
 QY 781 CAGCATGTGGGCCCCAGGCCGGGGACTCTGTAGCCACTGCTGGAGAGCCACTCCCATCT 840
 |||||
 Db 1536 CAGCATGTG--CCCAGGCCGGGGACTCTGTAGCCACTGCTGGAGA--CCACTCCCATCT 1592
 |||||
 QY 841 TTCTCCCAATTTCTTACAGCTGCTGCCT----- 868
 |||||
 Db 1593 TTCTCCCAATTTCTTACAGCTGCTGCCTTATACAGTCACTTTT-----TTTGTAGACGG 1652
 |||||
 QY 869 ----- 868
 |||||
 Db 1653 AGTCTCGCTGTGTGCCCCAGGCTGGAGTGCAGTGGCGGGATCTCGGCTCACTGCAACGTC 1712
 |||||
 QY 869 ----- 868
 |||||

Db 589 CCAGCCCGTTTGAAGAGCTCACCCCGACACAAGGACCCGCACACAGATACCTCCAGC 648
 QY 61 TCCCTCTCAACCCACCCCTTTCCAGGTTGAGAACTTGGAGCAATAACTTGTCTCCATGA 120
 Db 649 TCCCTCTCAACCCACCCCTTTCCAGGTTGAGAACTTGGAGCAATAACTTGTCTCCATGA 708
 QY 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCACTTAGAACAA 180
 Db 709 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCACTTAGAACAA 768
 QY 181 TGACAAATAGAGGGGAATGGAATAAACAAGAGAAACGGTTTCCAGGACAGGGTT 240
 Db 769 TGACAAATAGAGGGGAATGGAATAAACAAGAGAAACGGTTTCCAGGACAGGGTT 828
 QY 241 TGGCTTACAGTTGTGGATGTGGGTACCCATGCCATGCCAAGTGTGAGGGAGGCTGCCGGTG 300
 Db 829 TGGCTTACAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGCCGGTG 888
 QY 301 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGAG 360
 Db 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGAG 947
 QY 361 GCCGGAGTTGAGACAGCTGGCCCAACATGCTGAAACCCCATCTGTACTAAAAATACA 420
 Db 948 GCCGGAGTTGAGACAGCTGGCCCAACATGCTGAAACCCCATCTGTACTAAAAATACA 1007
 QY 421 AAGTTAGCTGGCGGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 480
 Db 1008 AAGTTAGCTGGCGGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 1067
 QY 481 TGAGAAATCGCTTGAGCCAGCTGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 540
 Db 1068 TGAGAAATCGCTTGAGCCAGCTGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 1127
 QY 541 ACAGAAATAGTTGAGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 600
 Db 1128 ACAGAAATAGTTGAGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 1187
 QY 601 GATGAAGATTGCTTGAGCCTGGAGGTCAAGGCTGCAAGTGGAGGCTGAGGCTGAGGCA 660
 Db 1188 GATGAAGATTGCTTGAGCCTGGAGGTCAAGGCTGCAAGTGGAGGCTGAGGCTGAGGCA 1247
 QY 661 GCATCCAGCTGGGCAACAGAGTGAGACCTGTCTCGAAAAAATAAATAAATAAATAA 720
 Db 1248 GCATCCAGCTGGGCAACAGAGTGAGACCTGTCTCGAAAAAATAAATAAATAAATAA 1305
 QY 721 GGAGAGGAGAGACTCAAGCAGCCCTTCAAGAGCTGCTGAGGCTGAGGCTGAGGCTG 780
 Db 1306 GGAGAGGAGAGAGACTCAAGCAGCCCTTCAAGAGCTGCTGAGGCTGAGGCTGAGGCTG 1365
 QY 781 CAGCATGTGGCCCCAGGCGGGACTCTGTAGCCACTGCTGAGAGCCACTTCCCATCT 840
 Db 1366 CAGCATGTG--CCAGGCGGGGACTCTGTAGCCACTGCTGAGAGCCACTTCCCATCT 891
 QY 841 TTCTCCCATTTCTAGACCTGCTGCTATACAGTCACTTTTATGTGTTTT 891
 Db 1423 TTCTCCCATTTCTAGACCTGCTGCTATACAGTCACTTTTATGTGTTTT 1473

RESULT 4

ADF39358

ID ADF39358 standard; DNA; 1856 BP.

XX

AC ADF39358;

XX

DT 12-FEB-2004 (first entry)

XX

Intron 16 sequence of human ACE gene.

DE

KW Diabetes mellitus; renin-angiotensin II system;

KW angiotensin-converting enzyme; ACE; antidiabetic; insulin; hypoglycaemia;

KW ACE genotype; intron 16; human; ds.

XX

OS Homo sapiens.

XX US2003158090-A1.

XX 21-AUG-2003.

XX

XX 04-OCT-2002; 2002US-00195330.

XX

XX 23-JUL-2001; 2001US-0306859P.

XX

XX (PEDE/) PEDERSEN-BJERGAARD U.

XX (AGER/) AGERHOLM-LARSEN B.

XX (THOR/) THORSTEINSSON B.

XX (PRAM/) PRAMMING S.

XX

XX Pedersen-Bjergaard U, Agerholm-Larsen B, Thorsteinsson B;

XX Prammig S;

XX WPI; 2003-778014/73.

XX

XX Treating diabetes mellitus in an individual comprises administering an

XX inhibitor of the renin-angiotensin II system and an antidiabetic.

XX Claim 36; Page 13-14; 17pp; English.

XX

The present invention relates to a method of treating diabetes mellitus in an individual. The method comprises administering an inhibitor of the renin-angiotensin II system (e.g. angiotensin-converting enzyme (ACE)) and an antidiabetic (e.g. insulin). Also disclosed is a method of diagnosing susceptibility of an individual to hypoglycaemia. The method involves detecting within a tissue sample the genotype of the ACE gene and correlating the genotype with the susceptibility to hypoglycaemia. The methods are useful for treating diabetes mellitus and hypoglycaemia. The present sequence represents intron 16 of the human ACE gene.

Sequence 1856 BP; 403 A; 546 C; 483 G; 424 T; 0 U; 0 Other;

Query Match

Best Local Similarity 41.0%; Score 824.6; DB 10; Length 1856;

Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;

QY 1 CCGGCCCGTTTGAAGAGCTCACCCCGACACAAGGACCCGCACACAGATACCTCCAGC 60
 Db 589 CCAGCCCGTTTGAAGAGCTCACCCCGACACAAGGACCCGCACACAGATACCTCCAGC 648
 QY 61 TCCCTCTCAACCCACCCCTTTCCAGGTTGGAGAACTTGAGGCATAAACTTGTCTCATGA 120
 Db 649 TCCCTCTCAACCCACCCCTTTCCAGGTTGGAGAACTTGAGGCATAAACTTGTCTCATGA 708
 QY 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCACTTAGAACAA 180
 Db 709 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCACTTAGAACAA 768
 QY 181 TGACAAATAGAGGGGAATGGAATAAACAAGAGAAACGGTTTCCAGGACAGGGTT 240
 Db 769 TGACAAATAGAGGGGAATGGAATAAACAAGAGAAACGGTTTCCAGGACAGGGTT 828
 QY 241 TGGCCTCAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGCCGGTG 300
 Db 829 TGGCCTCAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGCCGGTG 888
 QY 301 TGGTGGCTCATGCTTAAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGA 360
 Db 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGA 947
 QY 361 GCCGGAGTTGAGACAGCTGGCCCAACATGCTGAAACCCCATCTGTACTAAAAATACA 420
 Db 948 GCCGGAGTTGAGACAGCTGGCCCAACATGCTGAAACCCCATCTGTACTAAAAATACA 1007
 QY 421 AAGTTAGCTGGCGGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 480
 Db 1008 AAGTTAGCTGGCGGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 1067

XX OS Homo sapiens.
 XX PN WO2003054166-A2.
 XX PD 03-JUL-2003.
 XX PF 19-DEC-2002; 2002WO-US041225.
 XX PR 20-DEC-2001; 2001US-0342603P.
 XX PA (INCY-) INCYTE GENOMICS INC.
 XX PI Jones KA, Schafer A;
 XX DR WPI; 2003-559141/52.
 XX KW Determining susceptibility of an individual to joint space narrowing,
 XX PT osteophyte development and/or joint pain comprises identifying whether
 XX PT the individual has at least one polymorphism in a polynucleotide encoding
 XX PT a protein.
 XX PS Disclosure; SEQ ID NO 54; 297pp; English.
 XX CC The invention relates to a method of determining susceptibility of an
 XX CC individual to joint space narrowing and/or osteophyte development and/or
 XX CC joint pain comprising identifying whether the individual has at least one
 XX CC polymorphism in a polynucleotide encoding at least one of the protein
 XX CC listed in the specification. The methods, composition and agent are
 XX CC useful for modulating the susceptibility of an individual to joint space
 XX CC narrowing and/or osteophyte development and/or joint pain that is
 XX CC associated with a disease, preferably osteoarthritis. The cell line and
 XX CC the non-human animal are useful for screening for an agent for diagnosing
 XX CC an individual having susceptibility to joint space narrowing and/or
 XX CC osteophyte development and/or joint pain. This sequence corresponds to
 XX CC the polynucleotide encoding a protein listed in the specification. (Note:
 XX CC The sequence data for this patent did not form part of the printed
 XX CC specification but was obtained in electronic format directly from WIPO at
 XX CC ftp.wipo.int/pub/published_pct_sequences).
 XX SQ Sequence 41369 BP; 9847 A; 10483 C; 11706 G; 9333 T; 0 U; 0 Other;
 XX
 XX Query Match 13.4%; Score 270.4; DB 10; Length 41369;
 XX Best Local Similarity 77.1%; Pred. No. 6.5e-59;
 XX Matches 347; Conservative 0; Mismatches 91; Indels 12; Gaps 1;
 XX
 XX QY 292 GCGCGGTGTGGTCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGTGAGTAGA 351
 XX DB 33439 GGTGAGGCGATGGTGGTCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGTGAGTAGA 33498
 XX QY 352 TCACCTTGAGGCGCGGAGTTTGAGACGAGCTGCGCAACATGTGAACCCCATCTGTACT 411
 XX DB 33499 TCACCTGAGTACAGAGTTTGAGACGAGCTGCGCAACATGTGAACCCCATCTGTACT 33558
 XX QY 412 AAAAATACAAAGTTAGTGGCGGTGGTGTAGATGCCCTGTAGTCCAGCTACTTTGGGAG 471
 XX DB 33559 AAAAATACAAAGTTAGTGGCGGTGGTGTAGATGCCCTGTAGTCCAGCTACTTTGGGAG 33618
 XX QY 472 GCTGAGGCGATGAGATC-----GCTTGAGCCAGCTGGGCAATACAGACAGA 519
 XX DB 33619 GCGGAGGCGGGGATACAGAGATCAGGAGATTTAGACCATCTCTGGCTAACACAGTGAAA 33678
 XX QY 520 CCCCCTCTCTACAAATAAAATACAAATAATTTGGATGTGGTGTAGTCCAGCTGGGAGCTGCA 639
 XX DB 33679 CCTGTCTCTACAAATAAAATACAAATAATTTAGCCGGCGGTAGTGGCGGCGCTGTAGT 33738
 XX QY 580 CTTAGCTGCTAGGAGGCTGAGATGGAAGGATTTGCTTGAGCTGGGAGCTCAAGCTGCA 639
 XX DB 33739 CCGAGCTATTTCGGAGGCTGAGCAGGAGATGGGCTGNAACCGAGAGAGGAGGCTTGA 33798
 XX QY 640 GTGACCCAGATGGGCGCACTGCACTCAGCCTGGGCAACAGAGTGAGACCTGTCTCAG 699
 XX DB 33799 GTGACCCAGATGGGCGCACTGCACTCAGCCTGGGCAACAGAGTGAGACCTGTCTCAG 33858

QY 700 AAAAAAAAAAAAAAAAAAAAAAAAAAGGAGGAG 729
 DB 33859 AAAAAATATAAAATATAAAATATAAAATAG 33888
 RESULT 7
 ADM01345/C
 ID ADM01345 standard; cDNA; 2836 BP.
 XX
 AC ADM01345;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human cDNA of the invention SEQ ID NO:30.
 XX
 KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
 XX
 OS Homo sapiens.
 XX
 PN EP1347046-A1.
 XX
 PD 24-SEP-2003.
 XX
 PF 12-APR-2002; 2002EP-00008400.
 XX
 PR 22-MAR-2002; 2002JP-00137785.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 DR WPI; 2003-723558/69.
 XX
 DR P-PSDB; ADM03788.
 XX
 PT New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.
 XX
 PS Claim 1; SEQ ID NO 30; 305pp; English.
 XX
 CC The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC cDNA sequence of the invention.
 XX
 SQ Sequence 2836 BP; 632 A; 722 C; 676 G; 806 T; 0 U; 0 Other;
 XX
 XX Query Match 13.3%; Score 267.2; DB 11; Length 2836;
 XX Best Local Similarity 78.8%; Pred. No. 1.5e-58;
 XX Matches 350; Conservative 0; Mismatches 78; Indels 16; Gaps 2;
 XX
 XX QY 292 GCGCGGTGTGGTCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGTGAGTAGA 351
 XX DB 2473 GGCTGGGCACAGTGGCTCACACCTGTATATCCAGCACTCTAGGAGGCCAAGTGAGTAGA 2414
 XX QY 352 TCACCTTGAGGCGCGGAGTTTGAGACGAGCTGCGCAACATGTGAACCCCATCTGTACT 411
 XX DB 2413 CCACCTTGAGTACAGAGTTTCAGACCCAGCTGGCCACACCTCGTCTCTACT 2354
 XX QY 412 AAAAATACAAAGTTAGTGGCGGTGGTGTAGATGCCCTGTAGTCCAGCTACTTTGGGAG 471
 XX DB 2353 GAAAAATATAAAATTAGCCAGGATGGTGGCACAGCTGTATATCCAGCAAGTTGGGAG 2294
 XX QY 472 GCTGAGGCGATGAGATC-----GCTTGAGCCCGAGCTGGGCAATACAGCAA 517


```
XX 01-MAY-2003.
XX 22-OCT-2002; 2002WO-US033835.
XX 23-OCT-2001; 2001US-00004113.
XX (SAGR-) SAGRES DISCOVERY.
XX Engelhard EK, Morris DW;
XX WPI; 2003-421412/39.
XX New recombinant nucleic acid and its encoded protein, useful for
XX preparing a composition for diagnosing or treating carcinomas.
XX Claim 1; Page 103-107; 173pp; English.
XX The invention relates to novel sequences which are useful for preparing a
XX composition for diagnosing or treating carcinomas. These sequence are
XX also useful in gene therapy. The present sequence is human CCND3
XX carcinoma associated (CA) gene. This sequence is used in the invention
XX
XX Sequence 26874 BP; 6621 A; 6603 C; 6750 G; 6880 T; 0 U; 20 Other;
XX
XX Query Match 13.1%; Score 262.8; DB 8; Length 26874;
XX Best Local Similarity 76.6%; Pred. No. 5.1e-57;
XX Matches 340; Conservative 0; Mismatches 92; Indels 12; Gaps 1;
XX
XX QY 292 GCGCGGGTGGTGCTCATGCTTAATCCAGCAGCTTTGGAGGCGCAAGTGAGTAGA 351
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 5012 GCTGGGGCGGTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCGCGAGTGGCGGA 4953
XX
XX QY 352 TCACCTTGAGCGCGGAGTTTGAGACAGCTGGCCCAACATGGTGAACCCCATCTGTACT 411
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4952 TCACCTGAAGTCAGAGGTTCAAGACAGCTGGCCCAACATGGCAAAACCCCATCTTACT 4893
XX
XX QY 412 AAAAAATACAAAAGTTAGTGGCGGTGGTGGTAGATGCTCTAGTCCAGCTACTCTGGGAG 471
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4892 AAAAAATACAAAAGTTGGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4833
XX
XX QY 472 GCTGAGGCGATGAGATC-----GTTGAGCGCGAGCTGGGCAATACAGCAAGA 519
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4832 GCGGAGGCGGCGAGATCAGAGTCAGGAGATCGAGACAGCTTAACCAACATGTTGAAA 4773
XX
XX QY 520 CCCCCTCTCTACAAATACAAAATTAAGTTGGATGTTGGTGGTGGTGGTGGTGGTGGT 579
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4772 CCCCCTCTCTCTACAAATACAAAATTAAGTTGGGCGATGGTGGGCGATGCCCTGTAA 4713
XX
XX QY 580 CCTAGCTGTAGGAGGCTGAGATGGAAGGATTTGCTTGAGCCTGGGAGGTCAAGGCTGCA 639
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4712 CCAGCTACTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4653
XX
XX QY 640 GTGAGCGGAGATGGCGGCTCTGCACTCCAGCTGGGCAACAGAGTGAGACCTGTCTGAG 699
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4652 GTGAGCTGAGATCAGCGGCTGCACTCCAGCTGGGCAACAGAGTGAGACCTCGTCTCAA 4593
XX
XX QY 700 AAAAAAATAAAAAAATAAAAAAAGGA 723
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4592 AAAAAAATAAAAAAATAAAAAAAGGA 4569
XX
XX RESULT 12
XX ADA02467/c
XX ID ADA02467 standard; DNA; 26874 BP.
XX AC ADA02467;
XX AT
XX DT 06-NOV-2003 (first entry)
XX DE Human CCND3 carcinoma associated gene, SEQ ID NO:986.
XX KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
```

```
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX Homo sapiens.
XX WO2003057146-A2.
XX PD 17-JUL-2003.
XX PF 26-DEC-2002; 2002WO-US041414.
XX PR 26-DEC-2001; 2001US-00035832.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX DR WPI; 2003-587068/55.
XX PT New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.
XX PS Claim 1; SEQ ID NO 986; 245pp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically
XX binds to the protein, and a biochip comprising CA nucleic acid or
XX fragments thereof. The sequences of the invention were identified using
XX oncogenic retroviruses, which insert into the genome of the host organism
XX at random. Many of these do not carry transduced host oncogenes or
XX pathogenic trans-acting viral genes, meaning that cancer incidence is a
XX direct consequence of the effects of proviral integration into host
XX protooncogenes. The CA nucleic acid sequences can be used to diagnose
XX carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX leukaemia) or a propensity to carcinoma by determination of the sequence
XX of a CA gene, or by determination of CA gene expression in particular
XX tissues. CA nucleic acids, proteins and antibodies are also useful as
XX therapeutic agents and in screening and evaluating drug candidates. The
XX present sequence represents a specifically claimed human CA nucleic acid
XX sequence of the invention. Note: The complete sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 26874 BP; 5621 A; 6602 C; 6751 G; 6880 T; 0 U; 20 Other;
XX
XX Query Match 13.1%; Score 262.8; DB 9; Length 26874;
XX Best Local Similarity 76.6%; Pred. No. 5.1e-57;
XX Matches 340; Conservative 0; Mismatches 92; Indels 12; Gaps 1;
XX
XX QY 292 GCGCGGGTGGTGCTCATGCTTAATCCAGCAGCTTTGGAGGCGCAAGTGAGTAGA 351
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 5012 GCTGGGGCGGTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCGCGAGTGGCGGA 4953
XX
XX QY 352 TCACCTTGAGCGCGGAGTTTGAGACAGCTGGCCCAACATGGTGAACCCCATCTGTACT 411
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4952 TCACCTGAAGTCAGAGGTTCAAGACAGCTGGCCCAACATGGCAAAACCCCATCTTACT 4893
XX
XX QY 412 AAAAAATACAAAAGTTAGTGGCGGTGGTGGTAGATGCTCTAGTCCAGCTACTCTGGGAG 471
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4892 AAAAAATACAAAAGTTGGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4833
XX
XX QY 472 GCTGAGGCGATGAGATC-----GTTGAGCGCGAGCTGGGCAATACAGCAAGA 519
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4832 GCGGAGGCGGCGAGATCAGAGTCAGGAGATCGAGACAGCTTAACCAACATGTTGAAA 4773
XX
XX QY 520 CCCCCTCTCTACAAATACAAAATTAAGTTGGATGTTGGTGGTGGTGGTGGTGGTGGT 579
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4772 CCCCCTCTCTCTACAAATACAAAATTAAGTTGGGCGATGGTGGGCGATGCCCTGTAA 4713
XX
```



```
FT /number= 4
FT intron 4919. .5702
FT /*tag= i
FT /number= 4
FT variation replace(5539,G)
FT /*tag= ae
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(5558,G)
FT /*tag= af
FT /standard_name= "Single nucleotide polymorphism"
FT exon 5703. .5853
FT /*tag= j
FT /number= 5
FT intron 5854. .6056
FT /*tag= k
FT /number= 5
FT variation replace(5861,T)
FT /*tag= ag
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(6023,G)
FT /*tag= ah
FT /standard_name= "Single nucleotide polymorphism"
FT exon 6057. .6230
FT /*tag= l
FT /number= 6
FT intron 6231. .6389
FT /*tag= m
FT /number= 6
FT exon 6390. .6506
FT /*tag= n
FT /number= 7
FT intron 6507. .8832
FT /*tag= o
FT /number= 7
FT variation replace(6799,T)
FT /*tag= ai
FT /standard_name= "Single nucleotide polymorphism"
FT exon 8833. .9114
FT /*tag= p
FT /number= 8
FT intron 9115. .9885
FT /*tag= q
FT /number= 8
FT variation replace(9579,A)
FT /*tag= aj
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(9842,C)
FT /*tag= ak
FT /standard_name= "Single nucleotide polymorphism"
FT exon 9886. .9963
FT /*tag= r
FT /number= 9
FT intron 9964. .10201
FT /*tag= s
FT /number= 9
FT variation replace(10159,C)
FT /*tag= al
FT /standard_name= "Single nucleotide polymorphism"
FT exon 10202. .10324
FT /*tag= t
FT /number= 10
FT intron 10325. .10638
FT /*tag= u
FT /number= 10
FT exon 10639. .10754
FT /*tag= v
FT /number= 11
FT intron 10755. .15675
FT /*tag= w
FT /number= 11
FT variation replace(12025,G)
FT /*tag= am
FT /standard_name= "Single nucleotide polymorphism"
```

```
FT variation replace(14723,C)
FT /*tag= an
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14996,A)
FT /*tag= ao
FT /standard_name= "Single nucleotide polymorphism"
FT exon 15676. .15817
FT /*tag= x
FT /number= 12
FT intron 15818. .16071
FT /*tag= y
FT /number= 12
FT exon 16072. .16108
FT /*tag= z
FT /number= 13
FT intron 16109. .16828
FT /*tag= aa
FT /number= 13
FT variation replace(16153,G)
FT /*tag= ap
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(16181,A)
FT /*tag= ag
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(16756,G)
FT /*tag= ar
FT /standard_name= "Single nucleotide polymorphism"
FT exon 16829. .17008
FT /*tag= ab
FT /number= 14
FT intron 17009. .18491
FT /*tag= ac
FT /number= 14
FT variation replace(18059,G)
FT /*tag= as
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(18364,T)
FT /*tag= at
FT /standard_name= "Single nucleotide polymorphism"
FT exon 18492. .18568
FT /*tag= ad
FT /number= 15
FT variation replace(18861,A)
FT /*tag= au
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(20443,A)
FT /*tag= av
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(20881,T)
FT /*tag= aw
FT /standard_name= "Single nucleotide polymorphism"
FT XX WO200272765-A2.
FT PD 19-SEP-2002.
FT XX
FT PF 08-MAR-2002; 2002WO-US007159.
FT XX
FT PR 14-MAR-2001; 2001US-00805455.
FT XX
FT PA (PEXE ) PE CORP NY.
FT XX
FT PI Gan W, Ye J, Di Francesco V, Beasley EM;
FT DR WPI; 2002-750490/81.
FT DR P-PSDB; AAO26354.
FT XX
FT PT New isolated Ras-like protein polypeptides, useful for treating AIDS,
FT PT neurodegenerative diseases, ischemic injuries, toxin-induced diseases,
FT PT viral infections, cancer and osteoporosis.
FT XX
FT PS Claim 4; Fig 3A-G; 91pp; English.
FT XX
```


Db 935 TGCATGCTACGTGCGCGGCCCTGACCGTCACTACGGGGCCAGCACATCAACCTGG 994
Qy 1445 AGGGGGCCATTCTGCTCACCCTGCTGGGTAA 1475
Db 995 AGGGGGCCATTCTGCTCACCCTGCTGGGGAA 1025

Search completed: November 22, 2004, 17:23:08
Job time : 901.868 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:54:56 ; Search time 8217.69 Seconds
(without alignments)
12349.531 Million cell updates/sec

Title: US-10-017-724-1_COPY_3000_5784

Perfect score: 2785

Sequence: 1 gaggacttgacggtgatgg.....aattgtaaaaggtttct 2785

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	831	29.8	939	5	BQ933301
2	791.8	28.4	891	5	BQ925776
3	749.8	26.9	763	7	CK000113
4	749.2	26.9	937	5	BQ934054
5	715.4	25.7	757	6	CA449275
6	704.8	25.3	752	6	CD511356
7	704.2	25.3	746	5	BX114562
8	700.2	25.1	876	1	AU121101
9	694.8	24.9	721	4	BM718851
10	689.4	24.8	778	6	CA414667
11	685.8	24.6	707	7	CN274716
12	685.4	24.5	793	1	A1346330
13	682.8	24.5	748	4	BG680063
14	681.2	24.5	750	5	BUE18192
15	677.8	24.3	681	7	CN274719
16	675.8	24.3	756	5	BQ008487
17	663.2	23.8	706	2	BF109743
18	651.2	23.4	668	7	CN274711
19	649	23.3	693	2	BG673014
20	646.8	23.2	714	6	CA428685
21	645.2	23.2	660	1	A1419866
22	643	23.1	672	6	CD512572
23	638.8	22.9	660	6	CA431209
24	638.8	22.9	661	4	BM272400

C 25	635.8	22.8	673	5	BQ002730	UI-H-EI1-
C 26	632.6	22.7	706	2	AM190828	AW190828 x142C04.x
C 27	629.8	22.6	651	4	BG939270	CG939270 cn31e12.x
C 28	625.2	22.4	647	6	CA425254	CA425254 UI-H-DF0-
C 29	623.4	22.4	712	6	CA420699	CA420699 UI-H-FG0-
C 30	623.2	22.4	718	1	A1453687	A1453687 t339a07.x
C 31	622	22.3	648	6	CA421578	CA421578 UI-H-FG0-
C 32	620.2	22.3	558	1	AA149859	AA149859 z147e10.x
C 33	619.2	22.2	707	5	BUE27598	BUE27598 UI-H-FG0-
C 34	617.6	22.2	642	5	BQ019822	BQ019822 UI-H-ED0-
C 35	615.4	22.1	701	6	CA446830	CA446830 UI-H-ED1-
C 36	610	21.9	826	6	CD511191	CD511191 AGENCOURT
C 37	609.8	21.9	708	5	BM997135	BM997135 UI-H-ED0-
C 38	609	21.9	817	4	B1523839	B1523839 603051961
C 39	608.8	21.9	637	1	AA625208	AA625208 af70f01.x
C 40	608.8	21.9	660	5	BM988626	BM988626 UI-H-DH0-
C 41	607.8	21.8	3371	9	AY398790	AY398790 Homo sapi
C 42	606.8	21.8	610	7	CN481711	CN481711 hwl1c03.y
C 43	604.2	21.7	659	6	CA422281	CA422281 UI-H-FG0-
C 44	600.4	21.6	655	5	BM971657	BM971657 UI-CF-EC1
C 45	600	21.5	648	1	AI571506	AI571506 tr85a06.x

ALIGNMENTS

RESULT 1
BQ933301
LOCUS
DEFINITION
AGENCOURT 8744271 Lupski sciatic_nerve Homo sapiens CDNA clone
IMAGE:6205936 5', mRNA sequence.
ACCESSION
BQ933301
VERSION
BQ933301.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 939)
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupeki
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MSC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13629 row: j column: 17
High quality sequence stop: 686.

FEATURES
Location/Qualifiers
1...939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205936"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
Not1: Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCTCCG-3' and
5'-GACTAGTCTAGATCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life

ORIGIN		Technologies."	
Query Match		29.8%; Score 831; DB 5; Length 939;	
Best Local Similarity		96.9%; Pred. No. 3.4e-189;	
Matches 912; Conservative 0; Mismatches 17; Indels 12; Gaps 6;			
Qy	1815	CTAGTAACACATAGTAGTATAATTTCTAGAAAACAAACATCTAATAAGTATATATATCTG	1874
Db	1	CTAGTAACACATAGTAGTATAATTTCTAGAAAACAAACATCTAATAAGTATATATCTG	60
Qy	1875	TGAAAATATAGGGCTTGATAATATATAGTTCTCAGATGAAGCATGTAGAGCTGAAC	1934
Db	61	TGAAAATATAGGGCTTGATAATATATAGTTCTCAGATGAAGCATGTAGAGCTGAAC	120
Qy	1935	AGAAATACATAGAGAAATATAGAGGATTTATGATGGAACCTTAATATATATGTTGCCAGC	1994
Db	121	AGAAATACATAGAGAAATATAGAGGATTTATGATGGAACCTTAATATATATGTTGCCAGC	180
Qy	1995	GATTTTACTCAATATTTGTTTACTGTTATCTATCTGCTGTATATGGAATCTTTTAATTC	2054
Db	181	GATTTTACTCAATATTTGTTTACTGTTATCTATCTGCTGTATATGGAATCTTTTAATTC	240
Qy	2055	AAACGCTGAAAACGAATCAGCATTTAGTCTTGGCAGGCACACCCCAATATCAGTCATGTG	2114
Db	241	AAACGCTGAAAACGAATCAGCATTTAGTCTTGGCAGGCACACCCCAATATCAGTCATGTG	300
Qy	2115	TAATATGCAAGTTGTTTGTGTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2170
Db	301	TAATATGCAAGTTGTTTGTGTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	360
Qy	2171	TTAAGTCTCATGATCTTTCTGCAGAAATAGTCACTCATCCACATCCACATAAGGGTTT	2230
Db	361	TTAAGTCTCATGATCTTTCTGCAGAAATAGTCACTCATCCACATCCACATAAGGGTTT	420
Qy	2231	AGTAAGAGAAGTCTGCTGTCATGATGATGAGGAGGGAATCTTTTCCCTTTCTGT	2290
Db	421	AGTAAGAGAAGTCTGCTGTCATGATGATGAGGAGGGAATCTTTTCCCTTTCTGT	480
Qy	2291	TAATAGTCTACATCTTTCTATGCCAAACAGGAACCATCATCACTTATGCTTATATGATAC	2350
Db	481	TAATAGTCTACATCTTTCTATGCCAAACAGGAACCATCATCACTTATGCTTATATGATAC	540
Qy	2351	ACATTTGCTTTTATGATAAATTAATTTGTTGTTTCTTTTGGAGTTGATGTTGTTGTTGTTG	2407
Db	541	ACATTTGCTTTTATGATAAATTAATTTGTTGTTTCTTTTGGAGTTGATGTTGTTGTTGTTG	600
Qy	2408	GTTTGTCTGCACTTTTACTTTTTCGCTGCTGAGCTGATTTCCCGAGA-CAACGAAGCG	2466
Db	601	GTTTGTCTGCACTTTTACTTTTTCGCTGCTGAGCTGATTTCCCGAGA-CAACGAAGCG	660
Qy	2467	TTGGGATCTTCAATTAATGATGAGTCTGCAACGCGTGCAGGTTTCTGTTCTGTTGTTG	2526
Db	661	TTGGGATCTTCAATTAATGATGAGTCTGCAACGCGTGCAGGTTTCTGTTCTGTTGTTG	720
Qy	2527	TGTGGGCTCAACCGTACAAATGTTGGGAATGACGATGATGAAATATTAGAAATGATACC	2586
Db	721	TGTGGGCTCAACCGTACAAATGTTGGGAATGACGATGATGAAATATTAGAAATGATACC	780
Qy	2587	ATATTTTTTTGTAATTTATTTATGTTTCTAAACAAATTTATGTTATGTTGATGAAAC	2646
Db	781	ATATTTTTTTG-AAAATATTATGTTTCTAA-CAATTTATGTTATGTTGATGAAAC	837
Qy	2647	GTCATGTTTGTCCAAAGACTGTAATATTTATTTATGTTGTTGTTGTTGTTGTTGTTGTTG	2706
Db	838	GTCATGTTTGTCCAAAGACTGGAATAATTTATTTATGTTGTTGTTGTTGTTGTTGTTG	897
Qy	2707	ACCAGTCTGCACTTATGCTAGAAACCTTCAATTTTAA	2746
Db	898	ACCAGTCTGCACTTATGCTAGAAACCTTCAATTTTAA	938

```
QY 1451 ACAACATCTTGCATAGGTTGAGCGGGTTCAGATGTGGATTGGCAAACTCAT 1510
Db |||||||
QY 301 ACAACATCTTGCATAGGTTGAGCGGGTTCAGATGTGGATTGGCAAACTCAT 360
Db |||||||
QY 1511 TTAAGTAAAGGTTAGCAGCAAGTGGCGGTCTTTAGCTGTCTGTGCGGTGTGG 1570
Db |||||||
QY 361 TTAAGTAAAGGTTAGCAGCAAGTGGCGGTCTTTAGCTGTCTGTGCGGTGTGG 420
Db |||||||
QY 1571 CGTCGGGGAGGTCCTGCGCTGAGCTTCTTCCCGAGTTTGTGCTGAGAGCAACG 1630
Db |||||||
QY 421 CGTCGGGGAGGTCCTGCGCTGAGCTTCTTCCCGAGTTTGTGCTGAGAGCAACG 480
Db |||||||
QY 1631 GCAGACGCACAGCGCGGAAAGCGCATTAACGGCTATCTAGGCTTTGGTAACGCG 1690
Db |||||||
QY 481 GCAGACGCACAGCGCGGAAAGCGCATTAACGGCTATCTAGGCTTTGGTAACGCG 540
Db |||||||
QY 1691 CAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGTTCCAGTTTAAAT 1750
Db |||||||
QY 541 CAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGTTCCAGTTTAAAT 600
Db |||||||
QY 1751 GTTAATATTTTAAAGTACTATAGATCAATCACTTACAGTAACTTATTTAAAT 1810
Db |||||||
QY 601 GTTAATATTTTAAAGTACTATAGATCAATCACTTACAGTAACTTATTTAAAT 660
Db |||||||
QY 1811 ATGCTAGTAACACATATGATATAATTTCTAGAAACAAACATCTAATAAGTATAAT 1870
Db |||||||
QY 661 ATGCTAGTAACACATATGATATAATTTCTAGAAACAAACATCTAATAAGTATAAT 720
Db |||||||
QY 1871 CCTGTGAAATATAGGCTTGATATAATTTAGGTTGTCAAGTAAAGCATCTAGAAGCT 1929
Db |||||||
QY 721 CCTGTGAAATATAGGCTTGATATAATTTAGGTTGTCAAGTAAAGCATCTAGAAGCT 780
Db |||||||
QY 1930 GTAACAGATACATAGA-GAATAAGTGGAGTTATGATGGAACTTAAATATA 1983
Db |||||||
QY 781 GTAACAGATACATAGAATAGGAGGTTATGATGGAACTTAAATATA 835
Db |||||||
```

RESULT 3

```
CK000113
LOCUS CK000113
DEFINITION AGENCOURT_16368996 NIH_MGC_220 Homo sapiens cDNA clone
IMAGE:30707733 5', mRNA sequence.
```

```
ACCESSION CK000113
VERSION CK000113.1 GI:38526147
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

```
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
NIH-MGC http://mgi.nci.nih.gov/
```

```
1 (bases 1 to 763)
```

```
National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
Unpublished (1999)
```

```
Contact: Daniela S. Gerhard, Ph.D.
```

```
Office of Cancer Genomics
```

```
National Cancer Institute / NIH
```

```
Bldg. 31 Rm10A07 Bethesda, MD 20892
```

```
Email: cgapbs-r@mail.nih.gov
```

```
Tissue Procurement: James Martin, University of Iowa
```

```
cDNA Library Preparation: M. Bento Soares, University of Iowa
```

```
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
DNA Sequencing by: Agencourt Bioscience Corporation
```

```
Clone distribution: MGC clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
```

```
Plate: NDAM1072 row: m column: 22
```

```
High quality sequence stop: 664.
```

```
Location/Qualifiers
```

```
1..763
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="IMAGE:30707733"
```

FEATURES

source

```
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_220"
/notes="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
0.5-1kb. Adaptors 5'(AATTCGACGAGG)3' and 5'(
CTCTGTCGCG)3'. 3' Linker sequence - GCGGCGCTGAGACC T18.
Sequencing primers 3'end: T3 promoter primer 5'd
(ATTAACCTCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCACTATAGGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Average insert size 3-4kb
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match 26.9%; Score 749.8; DB 7; Length 763;
Best Local Similarity 99.5%; Pred. No. 1.3e-169;
Matches 751; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 66 AGATATTGATGATGTGTCTTGAACCAATGCCATCAGTGAGACAGACTTCAGGAACATT 125
Db |||||||
QY 8 AGATATTGATGATGTGTCTTGAACCAATGCCATCAGTGAGACAGACTTCAGGAACATT 67
Db |||||||
QY 126 CCAGATGTCCTTGGATCCCAAGGACCAACCAATTTGATCCAACTGGGTCAATTCG 185
Db |||||||
QY 68 CCAGATGTCCTTGGATCCCAAGGACCAACCAATTTGATCCAACTGGGTCAATTCG 127
Db |||||||
QY 186 CCATCAAGGCAAGGAGCTGGTTTCAGACAGCCAACTCGGACCCCGGCATCGTGTAGGTTT 245
Db |||||||
QY 128 CCATCAAGGCAAGGAGCTGGTTTCAGACAGCCAACTCGGACCCCGGCATCGTGTAGGTTT 187
Db |||||||
QY 246 TGACAGATTGGGTCTGTGGACTTCAGTGGGCACATTCTACGTAAACACTGACCGGACCA 305
Db |||||||
QY 188 TGACAGATTGGGTCTGTGGACTTCAGTGGGCACATTCTACGTAAACACTGACCGGACCA 247
Db |||||||
QY 306 CGACTATGCTGGCTTCGTCTTTGGTTTACAGTCAAGCAGCGCTTCTATGTGGTATGTG 365
Db |||||||
QY 248 CGACTATGCTGGCTTCGTCTTTGGTTTACAGTCAAGCAGCGCTTCTATGTGGTATGTG 307
Db |||||||
QY 366 GAAGCAGGTGACGACAGCTTACTGGGAGGACAGCCCAAGCGGGCTATGGCTACTCCGG 425
Db |||||||
QY 308 GAAGCAGGTGACGACAGCTTACTGGGAGGACAGCCCAAGCGGGCTATGGCTACTCCGG 367
Db |||||||
QY 426 CGTGTCTTCAAGGTGGTGAATCTCCACACGCGGAGCGGCGAGCACCTTGAGGAACCGCT 485
Db |||||||
QY 368 CGTGTCTTCAAGGTGGTGAATCTCCACACGCGGAGCGGCGAGCACCTTGAGGAACCGCT 427
Db |||||||
QY 486 GTGGCACAAGGGGAAACAGCGCGGGGAGGTGCGAACCTTTATGGCACGACCCAGGAACAT 545
Db |||||||
QY 428 GTGGCACAAGGGGAAACAGCGCGGGGAGGTGCGAACCTTTATGGCACGACCCAGGAACAT 487
Db |||||||
QY 546 TGGCTGGAGAGACTTACAGCGGCTTATAGTGGACCTGACTCAGAGGCCAAGACCGGCTA 605
Db |||||||
QY 488 TGGCTGGAGAGACTTACAGCGGCTTATAGTGGACCTGACTCAGAGGCCAAGACCGGCTA 547
Db |||||||
QY 606 CATCAGAGCTCTTAGTGATGAAGGAAACAGTTCATGTCAGACTCAGGACCTTCTATGA 665
Db |||||||
QY 548 CATCAGAGCTCTTAGTGATGAAGGAAACAGTTCATGTCAGACTCAGGACCTTCTATGA 607
Db |||||||
QY 666 CCAAACTTACGCTGGCGGGCTGGGTCTATTGTCTTCTCTCAAGAAATGGTCTATTT 725
Db |||||||
QY 608 CCAAACTTACGCTGGCGGGCTGGGTCTATTGTCTTCTCTCAAGAAATGGTCTATTT 667
Db |||||||
QY 726 CTCAGACCTCAAGTACGAATCAGAGATATTATAAACAAGATTTCGTCATTTCCGGCAAT 785
Db |||||||
QY 668 CTCAGACCTCAAGTACGAATCAGAGATATTATAAACAAGATTTCGTCATTTCCGGCAAT 727
Db |||||||
QY 786 GCCCTGTGCATGCCATGTCCTCCCTAGACACCTCAGT 820
Db |||||||
```

728 GCNCTGTGATGCATCGGTGGTCCCTAGACACCTCAGT 762

RESULT 4
 BQ934054
 LOCUS
 DEFINITION BQ934054 937 bp mRNA linear EST 21-AUG-2002
 AGENCOURT_8817273 Lupski_sciatic_nerve Homo sapiens cDNA clone
 IMAGE:6202836 5', mRNA sequence.
 BQ934054
 BQ934054.1 GI:22349437
 EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 937)
 NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LMAM3621 row: i column: 13
 High quality sequence stop: 613.
 Location/Qualifiers
 1. 937
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6202836"
 /sex="male"
 /tissue_type="sciatic nerve"
 /dev_stage="adult, 70 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski sciatic nerve"
 /notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
 NotI; Site 2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCACCGCGTCG-3' and
 5'-GACTACTGTAGTCAGCGCGGCCCT(15)-3'. Size selected >
 1 kb for average insert length 1.87 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

Query Match 26.9%; Score 749.2; DB 5; Length 937;
 Best Local Similarity 99.0%; Pred. No. 1.9e-169;
 Matches 754; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1126 AATACTGTGCATAGCCATTATGATAAATTAAGCATGAAAAAATTTGCTGAACCTACCTTT 1185
 DB 1 AATACTGTGCATAGCCATTATGATAAATTAAGCATGAAAAAATTTGCTGAACCTACCTTT 60

QY 1186 GGTGCTTAAAGTTGTCACCTATTCTTGAAATTAGAGTTGCTCTACAAATGACACACAAATCCC 1245
 DB 61 GGTGCTTAAAGTTGTCACCTATTCTTGAAATTAGAGTTGCTCTACAAATGACACACAAATCCC 120

QY 1246 GCTAAATAAATTTATAAACAAGGGTCAATTCAAAATTTGAAGTAATGTTTTAGTAAGGAGAG 1305
 DB 121 GCTAAATAAATTTATAAACAAGGGTCAATTCAAAATTTGAAGTAATGTTTTAGTAAGGAGAG 180

QY 1306 ATTAGAAGACAAACAGGCATAGCAATGACATATAGCTCCGATTAACCTAATTCGGAAACATGT 1365
 DB 181 ATTAGAAGACAAACAGGCATAGCAATGACATATAGCTCCGATTAACCTAATTCGGAAACATGT 240

	Query Match	25.3%;	Score 704.8;	DB 6;	Length 752;
	Best Local Similarity	98.9%;	Pred. No. 9.2e-159;		
	Matches 720;	Conservative 0;	Mismatches 7;	Indels 1;	Gaps 1;
QY	1591	GAGCTTCCTTCCCGCAGCTTTGCTGCTCAGAGGAAACGAGCAGCAGCAGCGCCGAAA	1650		
Db	3	GAGCTTCCTTCCCGCAGCTTTGCTGCTGAGAGGAAACGAGCAGCAGCGCCGAAA	62		
QY	1651	AGGCGCATCTAACCGGTATCTAGGCTTTGGTAACTCGCGACAAGTGTGCTTTTACCTGATT	1710		
Db	63	AGGCGCATCTAACCGGTATCTAGGCTTTGGTAACTCGCGACAAGTGTGCTTTTACCTGATT	122		
QY	1711	TGATGATACATTTTCATTAAGGTTCCAGTTATAATATTTTGGTTAATAATTTATTAAGTGAC	1770		
Db	123	TGATGATACATTTTCATTAAGGTTCCAGTTATAATATTTTGGTTAATAATTTATTAAGTGAC	182		

Db	301	TTCTATGTTGATGTGAAGCAGGTGACGAGACCTTCTGGAGGACACGCCACGCGG	360		
Qy	409	GCCTATGGCTACTCCGCGGTGTCCTCAAGTGTGTGAACCTCCACCGGGACGCGGCGAG	468		
Db	361	GCCTATGGCTACTCCGCGGTGTCCTCAAGTGTGTGAACCTCCACCGGGACGCGGCGAG	420		
Qy	469	CACCTGAGGAACCGCTGTGGCACACGGGGAACACCGCGGGGCGAGTGCAGACCTTATGG	528		
Db	421	CACCTGAGGAACCGCTGTGGCACACGGGGAACACCGCGGGGCGAGTGCAGACCTTATGG	480		
Qy	529	CACGACCCCAAGGAACATTGGCTGGAAGGACTACACGGCCTATAGTGGCACCCTGACTCAC	588		
Db	481	CACGACCCCAAGGAACATTGGCTGGAAGGACTACACGGCCTATAGTGGCACCCTGACTCAC	540		
Qy	589	AGGCCCCAAGACCGGCTACATCAGAGTCTTAGTGATGAAGGAAACACAGTCTATGGCGAC	648		
Db	541	AGGCCCCAAGACTGGCTACATCAGAGTCTTAGTGATGAAGGAAACACAGTCTATGGCGAC	600		
Qy	649	TCAGGACCTATCTATGACCAAACTACGCTGGCGGGCGGCTGCTATTGTGCTTCTCT	708		
Db	601	TCAGGACCTATCTATGACCAAACTACGCTGGCGGGCGGCTGCTATTGTGCTTCTCT	660		
Qy	709	CAAGAAATGCTTATTTCTCAGACCTCAAGTACCAATGCAGAGAT	753		
Db	661	CAAGAAATGCTTATTTCTCAGACCTCAAGTACCAATGCAGAGAT	705		
RESULT 9					
LOCUS	BW718851				
DEFINITION	721 bp mRNA linear EST 01-MAR-2002				
ACCESSION	UI-E-E01-ajc-j-23-0-UI.v1 UI-E-E01 Homo sapiens cDNA clone				
VERSION	UI-E-E01-ajc-j-23-0-UI 5', mRNA sequence.				
KEYWORDS	BW718851				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 721)				
TITLE	Bonaldo,M.F., Lennon,G. and Soares,M.B.				
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery				
MEDLINE	Genome Res. 6 (9), 791-806 (1996)				
FURNED	97044477				
COMMENT	8889548				
	Contact: Soares, MB				
	Coordinated Laboratory for Computational Genomics				
	University of Iowa				
	375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA				
	Tel: 319 335 8250				
	Fax: 319 335 9565				
	Email: bento-soares@uiowa.edu				
	Tissue procurement: Dr. Gregg Hageman				
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa				
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa				
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa				
	Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).				
	Seq primer: M13 Reverse.				
FEATURES	Location/Qualifiers				
source	1..721				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="UI-E-E01-ajc-j-23-0-UI"				
	/tissue_type="fetal eye"				
	/dev_stage="fetal"				
	/lab_host="DH10B (Life Technologies) (T1 phage resistant)"				
	/clone_lib="UI-E-E01"				
	/note="Organ: eye; Vector: pT7T3-Pac (pharmacia) with a modified polylinker. Site 1: EcoP I. Site 2: Not I."				

UI-E-EO1 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). "

ORIGIN

Query Match	24.9%;	Score 694.8;	DB 4;	Length 721;
Best Local Similarity	99.4%;	Pred. No. 2.4e-156;		
Matches 718;	Conservative 0;	Mismatches 2;	Indels 2;	Gaps 2;
Qy	922	GCCTTCAGAGGATAAATATCAATGGAACTCAGAGATGAACATCTAAACC-AC TAGAGGAA	980	
Db	1	GCCTTCAGAGGATAAATATCAATGGAAACGACAGATGAACATCTAAACCACATAGAGGAA	60	
Qy	981	ACCAGTTTGGTGATATATAGAGACTTTATGTGGAGTGAAAATTGGGGCATGCCATTACATTG	1040	
Db	61	ACCAGTTTGGTGATATATAGAGACTTTATGTGGAGTGAAAATTGGGGCATGCCATTACATTG	120	
Qy	1041	CTTTTCTTGTGTTTAAAAGAAATGACGTTTACATATAAAATGAAATTTACTTATTGTA	1100	
Db	121	CTTTTCTTGTGTTTAAAAGAAATGACGTTTACATATAAAATGAAATTTACTTATTGTA	180	
Qy	1101	TTTATGTGATATGGAGTTGCAAGGGAATCTGTGCATAAGCCATTATGATAAAATTAAACA	1160	
Db	181	TTTATGTGATATGGAGTTGCAAGGGAATCTGTGCATAAGCCATTATGATAAAATTAAACA	240	
Qy	1161	TGAAAATAATGTGCGTAACACTTTTGGTGCTTAAAGTTGTCACATATCTTGAATTAGAGT	1220	
Db	241	TGAAAATAATGTGCGTAACACTTTTGGTGCTTAAAGTTGTCACATATCTTGAATTAGAGT	300	
Qy	1221	TGCTCTACAAATGACACAAATCCGCTAAATAAATTTATAAAACAAGGTCAAATTCAAAT	1280	
Db	301	TGCTCTACAAATGACACAAATCCGCTAAATAAATTTATAAAACAAGGTCAAATTCAAAT	360	
Qy	1281	TGAAGTAATGTTTTAGTAGAGAGAGATTAGAAGACAAACAGGCATAGCAAAATGACATAAGC	1340	
Db	361	TGAAGTAATGTTTTAGTAGAGAGAGATTAGAAGACAAACAGGCATAGCAAAATGACATAAGC	420	
Qy	1341	TACCGATTAACATTCGGAACATGTAAAAACAGTTACAAAAATAAACGAATCTCCTCTTG	1400	
Db	421	TACCGATTAACATTCGGAACATGTAAAAACAGTTACAAAAATAAACGAATCTCCTCTTG	480	
Qy	1401	TCCTTACAAATGAAAGCCCTCATGTGCAGTAGAGATGCAGTTTTCATCAAGAAACAAACATCC	1460	
Db	481	TCCTTACAAATGAAAGCCCTCATGTGCAGTAGAGATGCAGTTTTCATCAAGAAACAAACATCC	540	
Qy	1461	TTGCAAAATGGGTGTGACCGCGTTCCAGATGTGGATTTGGCAAAACCTCATTTTAAGTAAAA	1520	
Db	541	TTGCAAAATGGGTGTGACCGCGTTCCAGATGTGGATTTGGCAAAACCTCATTTTAAGTAAAA	600	
Qy	1521	GGTTAGCAGAGCAAAAGTCGGTGCTTTAGCTGCTGCTTTGTGCCGTTGTGGCGTCGGGGAG	1580	
Db	601	GGTTAGCAGAGCAAAAGTCGGTGCTTTAGCTGCTGCTTTGTGCCGTTGTGGCGTCGGGGAG	660	
Qy	1581	GCTCCTGCTGAGCTTCCTTCCCGAGCTTTGCTGCTGAGAGGAACACGACAGACGCAC	1640	
Db	661	GCTCCTGCTGAGCTTCCTTCCCGAGCTTTGCTGCTGAGAGGAACACGACAGACGCAC	719	
Qy	1641	AG 1642		
Db	720	AG 721		

RESULT 10
CA414667/c

LOCUS
DEFINITION

ACCESSION

ACCESSION
VERSION
REFWORDC

KEYWORDS
SOURCE

ORGANISM

REFERENCE

[illegible]

JOURNAL.

COMMENT

FEATURES

source

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 793)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/Image.html
Insert Length: 1324 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 456.
Location/Qualifiers
1..793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1926422"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI-CCAP_Co8"
/note="Organ: colon; Vector: p7T73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo (dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified p7T73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

Query Match 24.5%; Score 683.4; DB 1; Length 793;
Best Local Similarity 95.8%; Pred. No. 1.4e-153;
Matches 744; Conservative 0; Mismatches 28; Indels 5; Gaps 4;

QY 393 GGACAGCCCGCGGCGCTATGGCTACTCCGGCGGTCCCTCAAGGTGGTGAATCCAC 452
DB 773 GGAGAACAGCCCGCGGCTATGGCTACTCCGGCGGTCCCTCAAGGTGGTGA 717
QY 453 CACGGGAGCGGCGAGCAGCCTGAGGAAACCGGCTGTGGCAGACAGCGGGGCA 512
DB 716 CACGGGAGCGGCGAGCAGCCTGAGGAAACCGGCTGTGGCAGACAGCGGGGCA 657
QY 513 GGTGGCACTTATGGCAGGACCCAGGAAACATTGGCTGGAGGACTACAGGCTATAG 572
DB 656 GGTGGCAA-CTTATGGCAGGACCCAGGAAACATTGGCTGGAGGACTACAGGCTATAG 598
QY 573 GTGGCACTGACTCAGCGCCCAAGCAGCGGTACATCAGAGTCTTAGTCATGAAGGAAA 632
DB 597 GTGGCACTGACTCAGCGCCCAAGCAGCGGTACATCAGAGTCTTAGTCATGAAGGAAA 538
QY 633 ACAGTCTATGGCAGACTCAGGACCTATCTATGACCAAACTACGCTGGCGGGCGGTGGG 692
DB 537 ACAGTCTATGGCAGACTCAGGACCTATCTATGACCAAACTACGCTGGCGGGCGGTGGG 478
QY 693 TCTATTGTCTCTCAGAGAAATGGTCTATTCTCAGACCTCAAGTACGATGACAGAGA 752
DB 477 TCTATTGTCTCTCAGAGAAATGGTCTATTCTCAGACCTCAAGTACGATGACAGAGA 418
QY 753 TATTTAAACAAGATTTGGTGCATTTCCGGCAATG-CCCTGTGCATGCCATGTGCCCTAGA 811
DB 417 TATTTAAACAAGATNTGCTGCATTTCCGGCAATGCCCCGTGTGATGCCATGTGCCCTAGA 358
QY 812 CACCTCAGTCAATGTGGTGGCGGCTTCTCTCTCTAGCAGCAGCCTCCTGTCCCTTGA 871

Db 357 CACCTCAGTTCATTTGGTCTCTCTCTAGCAGCAGCTCTCTGTCCCTTGA 298
QY 872 CTTTAACTCTGATGGTTCTTCCACCTCTGCAGCAACCCCAAAACCCCAAGTGCCTTCAGAG 931
DB 297 CTTTAACTCTGATGGTTCTTCCACCTCTGCAGCAACCCCAAAACCCCAAGTGCCTTCAGAG 238
QY 932 GATAAATATCAATGAAGAACTCAGAGATGAACATCTAACCCACTAGAGGAAACAGTTTGGT 991
DB 237 GATAAATATCAATGAAGAACTCAGAGATGAACATCTAACCCACTAGAGGAAACAGTTTGGT 178
QY 992 GATATATGAGACTTTATGTGGAGTGAATTTGGGCATGCCATTAACATTCCTTTCTTGT 1051
DB 177 GATATATGAGACTTTATGTGGAGTGAATTTGGGCATGCCATTAACATTCCTTTCTTGT 118
QY 1052 TTGTTTTAAAGAAATGACGTTTACATATAAAATTAATTAATTAATTAATTAATTAATTAAT 1111
DB 117 TTGTTTTAAAGAAATGACGTTTACATATAAAATTAATTAATTAATTAATTAATTAATTAAT 58
QY 1112 ATGAGTTGAAGGAATACGTGTGATAGCCATTAATTAATTAATTAATTAATTAATTAATTAAT 1168
DB 57 ATGAGTTGAAGGAATACGTGTGATAGCCATTAATTAATTAATTAATTAATTAATTAATTAAT 1
RESULT 13
LOCUS BG680063 748 bp mRNA linear EST 01-MAY-2001
DEFINITION 602628211P1 NCI-CCAP_Skn4 Homo sapiens cDNA clone IMAGE:4753130 5',
mRNA sequence.
ACCESSION BG680063 GI:13911460
VERSION BG680063.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 748)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10612 row: e column: 03
High quality sequence stop: 747.
Location/Qualifiers
1..748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4753130"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI-CCAP_Skn4"
/note="Organ: skin; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI-CCAP Library."

FEATURES
source

ORIGIN
Query Match 24.5%; Score 682.8; DB 4; Length 748;
Best Local Similarity 97.3%; Pred. No. 1.9e-153;
Matches 728; Conservative 0; Mismatches 12; Indels 8; Gaps 3;

QY 1720 ATTTTCATTAAGTTCAGTTTAAATATTTTCTTAATATTTTATTAAAGTACTATGAATG 1779
DB 1 ATTTTCATTAAGTTCAGTTTAAATATTTTCTTAATATTTTATTAAAGTACTATGAATG 60

```

QY 1780 CAACTCCATTACCAAGTAAGTATTTTAAATATGCGCTAGTAACACATATCTAGTATAATT 1839
Db 61 CAACTCCATTACCAAGTAAGTATTTTAAATATGCGCTAGTAACACATATCTAGTATAATT 120
QY 1840 TCTAGAAACAAACATCTAATAAGTATATAATCTGTGAAATATGAGGCTTGATAATTT 1899
Db 121 TCTAGAAACAAACATCTAATAAGTATATAATCTGTGAAATATGAGGCTTGATAATTT 180
QY 1900 AGTTGTGTCAGATGAAGCATGCTAGAGCTGTAAACAGATACATAGAGTATAGGAG 1959
Db 181 AGTTGTGTCAGATGAAGCATGCTAGAGCTGTAAACAGATACATAGAGTATAGGAG 240
QY 1960 TTTATGATGAAACCTTAAATATATAATAGTTGCCAGCGATTTTGTCTCAATATTTGTTACTG 2019
Db 241 TTTATGATGAAACCTTAAATATATAATAGTTGCCAGCGATTTTGTCTCAATATTTGTTACTG 300
QY 2020 TTATCTATCTGCTGTATATGGAATCTTTTAAATCAAAACGCTGAAAACGAATCAGCATTT 2079
Db 301 TTATCTATCTGCTGTATATGGAATCTTTTAAATCAAAACGCTGAAAACGAATCAGCATTT 360
QY 2080 AGTCTTGCCAGGCACACCAATCAATCAGTCATGTGTAATATGACACAGTTTGTGTTTGT 2139
Db 361 AGTCTTGCCAGGCACACCAATCAATCAGTCATGTGTAATATGACACAGTTTGTGTTTGT 420
QY 2140 TTGTTGTTTTTTT-----GTTGTTGTTGTTTTTTTAAAGTTGATGATCTTTCTGCGAG 2194
Db 421 TTGTTGTTTTTTTGTGTTGTTGTTTTTTTAAAGTTGATGATCTTTCTGCGAG 480
QY 2195 GAAATAGTCACTATCCCACTCCACATAAGGGGTTTAAAGAGAGTCTGCTGCTGA 2254
Db 481 GAAATAGTCACTATCCCACTCCACATAAGGGGTTTAAAGAGAGTCTGCTGCTGA 540
QY 2255 TCATGATAGGGGGCAAAATCTTTTCCCTTTCTGTTTAAATAGTATCATCATTTCTATGCC 2314
Db 541 TCATGATAGGGGGCAAAATCTTTTCCCTTTCTGTTTAAATAGTATCATCATTTCTATGCC 600
QY 2315 AAACAGGACGATCCATCACTTTAGTCTTAAATGACATTCATTCATTAATAAATTAAT 2374
Db 601 AAACAGGACGATCCATCACTTTAGTCTTAAATGACATTCATTCATTAATAAATTAAT 660
QY 2375 TTTGTTGTTTCTTTGAGGTTGATCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 2432
Db 661 TTTGTTGTTTCTTTGAGGTTGATCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 720
QY 2433 GCGTGTGGAGCTG-TATTCGCCAGACAA 2459
Db 721 GCGTGTGGAGCTGTTATTCGCCAGACAA 748

RESULT 14
BU618192/c
LOCUS
DEFINITION
UI-H-FH1-bfe-k-04-0-UI.s1 NCI CGAP FH1 Homo sapiens cDNA clone
UI-H-FH1-bfe-k-04-0-UI 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 750)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

```

Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu. The following repetitive elements were found in this cDNA sequence: 637-685, >(CAA)n#Simple_repeat
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FH1-bfe-k-04-0-UI"
/tissue_type="Cell Line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FH1"
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_FH1 is a normalized cDNA library obtained from a cell line derived from grade I chondrosarcoma tissue. The library was constructed and normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCCGCGC. The cell line was provided by Dr. James Martin from the University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
TAG_LIB=UI-H-FH1
TAG_SEQ=AGAATCCGCGC"

ORIGIN

Query Match 24.5%; Score 681.2; DB 5; Length 750;
Best Local Similarity 98.5%; Pred. No. 4.6e-153;
Matches 721; Conservative 0; Mismatches 3; Indels 8; Gaps 3;

QY 2062 GAAACAGGATCAGCATTTAGTCTTCCAGGACACCCATATCATGTGTAATATG 2121
Db 750 GAAACAGGATCAGCATTTAGTCTTCCAGGACACCCATATCATGTGTAATATG 691
QY 2122 CACAAGTTTGTGTTTTGTTTTGTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2177
Db 690 CACAAGTTTGTGTTTTGTTTTGTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 631
QY 2178 GCATGATCTTCTGAGGAAATAGTCACTCATCCACTCCACATAGGGGTTTACTAAGA 2237
Db 630 GCATGATCTTCTGAGGAAATAGTCACTCATCCACTCCACATAGGGGTTTACTAAGA 571
QY 2238 GAAGTCTGCTGCTGATGATGATAGGGGCAAACTTTTCCCTTCTGTTAATAGT 2297
Db 570 GAAGTCTGCTGCTGATGATGATGATAGGGGCAAACTTTTCCCTTCTGTTAATAGT 511
QY 2298 CATCACATTTCTATGCCAAACAGATCCATTAACCTTTAGTCTTATGTACACATGTC 2357
Db 510 CATCACATTTCTATGCCAAACAGATCCATTAACCTTTAGTCTTATGTACACATGTC 451
QY 2358 ATTTTGATAAAATTAATTTTGTGTTTCTTCTTGTAGGTTGATCGTTG---TGTTGTTTTC 2414
Db 450 ATTTTGATAAAATTAATTTTGTGTTTCTTCTTGTAGGTTGATCGTTGTTGTTGTTGTTGTT 391
QY 2415 TGCACCTTTTACTTTTTTGTGTTGAGCTGTATTCCGAGA-CAACGAACGTTGGAT 2473
Db 390 TGCACCTTTTACTTTTTTGTGTTGAGCTGTATTCCGAGACCAACGAACGTTGGAT 331
QY 2474 ACTTCATTAAATGATGAGCTGTCAACAGCTGAGGTTTCTGTTCTGTTGTTGTTGTTGTTGTT 2533
Db 330 ACTTCATTAAATGATGAGCTGTCAACAGCTGAGGTTTCTGTTCTGTTGTTGTTGTTGTTGTT 271

Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu. The following repetitive elements were found in this cDNA sequence: 637-685, >(CAA)n#Simple_repeat
Seq primer: M13 FORWARD
POLYA=Yes.

QY 2534 TCAACCGTACAAATGCTGTGGGATGACGATGATGTAATATTTAGAAATGTACCATATTTT 2593
 Db 270 TCAACCGTACAAATGCTGTGGGATGACGATGATGTAATATTTAGAAATGTACCATATTTT 211
 QY 2594 TTGTAAATTAATTTATGTTTTTCTAAACAAATTTATCGTATAGGTTGATGAAGAACTCATGT 2653
 Db 210 TTGTAAATTAATTTATGTTTTTCTAAACAAATTTATCGTATAGGTTGATGAAGAACTCATGT 151
 QY 2654 GTTTTGGCAAGACTGTAATATTTATTTATGTTTACATGTTTACATGTTTCAAAATTTACCACTG 2713
 Db 150 GTTTTGGCAAGACTGTAATATTTATTTATGTTTACATGTTTCAAAATTTACCACTG 91
 QY 2714 AAACCCCTGCACTAGCTAGCACTTCTTTTAAAGATTAAACACAGGAAATATAATTTGTAA 2773
 Db 90 AAACCCCTGCACTAGCTAGCACTTCTTTTAAAGATTAAACACAGGAAATATAATTTGTAA 31
 QY 2774 AAAAGGTTTTCT 2785
 Db 30 AAAAGGTTTTCT 19

RESULT 15
 CN274719
 LOCUS CN274719 681 bp mRNA linear EST 16-MAY-2004
 DEFINITION 17000532704744 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION CN274719
 VERSION CN274719.1 GI:47291133
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 681)
 AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
 Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
 Lebkowski,J. and Stanton,L.W.
 TITLE Transcriptome characterization elucidates signaling networks that
 control human ES cell growth and differentiation
 JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
 COMMENT Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@geron.com
 Insert Length: 681 Std Error: 0.00.

FEATURES
 source
 1..681
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, embryoid bodies
 derived from H1, H7 and H9 cells"
 /clone_lib="GRN EB"
 /note="oligo dt primed, full-length enriched cDNA library
 from embryoid body outgrowths derived from hES cell lines
 H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
 conditions."

ORIGIN

Query Match 24.3%; Score 677.8; DB 7; Length 681;
 Best Local Similarity 99.7%; Pred. No. 3e-152;
 Matches 679; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 129 GATGGTCCCCCTGGATCCCAAGGGACCAACCAAAATGATCCCAACTGGGTCAATTCGCCA 188
 Db 1 GATGGTCCCCCTGGATCCCAAGGGACCAACCAAAATGATCCCAACTGGGTCAATTCGCCA 60
 QY 189 TCAAGGCAAGGAGCTGGTTGACAGACCACTCGGACCCCGGCATCGCTGTAGGTTTGA 248
 Db 61 TCAAGGCAAGGAGCTGGTTGACAGACCACTCGGACCCCGGCATCGCTGTAGGTTTGA 120

Search completed: November 23, 2004, 12:26:12
 Job time : 8226.69 secs

QY 249 CGAGTTTGGGTCTGTGTGACCTTCAGTGGCAATTCTACGTAAACAACCTGACCGGGACGACGA 308
 Db 121 CGAGTTTGGGTCTGTGTGACCTTCAGTGGCAATTCTACGTAAACAACCTGACCGGGACGACGA 180
 QY 309 CTAATGCTGGCTTCGTCTTTGTTACCACTCAAGCAGCCGCTTCTATGTTGGTGTGTGAA 368
 Db 181 CTAATGCTGGCTTCGTCTTTGTTACCACTCAAGCAGCCGCTTCTATGTTGGTGTGTGAA 240
 QY 369 GCAGTGCAGCAGACCTACTTGGGAGGACCAAGCCCAACGCGGGCCCTATGGCTACTCCGGCGT 428
 Db 241 GCAGTGCAGCAGACCTACTTGGGAGGACCAAGCCCAACGCGGGCCCTATGGCTACTCCGGCGT 300
 QY 429 GTCCCTCAAGGTGTGTAACTCCACACGCGGAGCAGGCGAGCACTGAGGAAACGCGCTGTG 488
 Db 301 GTCCCTCAAGGTGTGTAACTCCACACGCGGAGCAGGCGAGCACTGAGGAAACGCGCTGTG 360
 QY 489 GCACACGGGACACGCGGGGAGGTGCGAACCTTATGGCAGCAGCCCGGAAACATTGG 548
 Db 361 GCACACGGGACACGCGGGGAGGTGCGAACCTTATGGCAGCAGCCCGGAAACATTGG 420
 QY 549 CTGGAAGGACTACACGGGCTATAGTGGCCTTACCTGACTCAGAGGCCCAAGACTGGCTACAT 608
 Db 421 CTGGAAGGACTACACGGGCTATAGTGGCCTTACCTGACTCAGAGGCCCAAGACTGGCTACAT 480
 QY 609 CAGAGTCTTATGTGATGAAGGAAACAGGTTCATGCGAGACTCAGGACCTATCTATGACCA 668
 Db 481 CAGAGTCTTATGTGATGAAGGAAACAGGTTCATGCGAGACTCAGGACCTATCTATGACCA 540
 QY 669 AACCTACGCTGGCGGGGCTGGGTCTATTTGTTCTCTCAAGAAATGGTCTATTTCTC 728
 Db 541 AACCTACGCTGGCGGGGCTGGGTCTATTTGTTCTCTCAAGAAATGGTCTATTTCTC 600
 QY 729 AGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTGCTGCAATTCGGCAATGCC 788
 Db 601 AGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTGCTGCAATTCGGCAATGCC 660
 QY 789 CTGTGATGCCATGGTCCCTA 809
 Db 661 CTGTGATGCCATGGTCCCTA 681

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:21 ; Search time 8322.11 Seconds
(without alignments)
11427.340 Million cell updates/sec

Title: US-10-017-724-3_COPY_85400_87410
Perfect score: 2011
Sequence: 1 ccggccggttgaaagagctt.....gaatgtggaagtgcgcgt 2011

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2011	100.0	98829	2	AC002345	AC002345 Homo sapi
c	2	2011	100.0	179509	9	AC113554 Homo sapi
	3	1997.4	99.3	23424	9	AY436326 Homo sapi
	4	1667.6	82.9	24070	9	AF118569 Homo sapi
	5	1233.6	61.3	2649	6	A28005 Homo sapi
	6	824.6	41.0	1856	6	AR000113 Homo sapi
	7	824.6	41.0	1856	6	AR137381 Sequence
	8	824.6	41.0	1856	6	BD211364 Genes for
	9	824.6	41.0	1856	6	BD075256 Methods f
	10	824.6	41.0	1856	6	BD075259 Methods f
	11	824.6	41.0	1856	9	HSATICE
	12	490.6	24.4	195863	2	AL596246 Mouse DNA
	13	299.8	14.9	144136	10	AC023518
c	14	299.8	14.9	217088	2	AC096435
c	15	277.2	13.8	213169	2	AC096705 Rattus no
	16	277.2	13.8	234451	2	AC096705 Rattus no
	17	273.8	13.6	188006	2	AC147422 Oryctolag
c	18	271.6	13.5	29488	9	AF503510 Homo sapi
	19	270.8	13.5	164824	2	AC023887 Homo sapi

20	270.6	13.5	172474	2	AC118582	AC118582 Oryctolag
21	270.4	13.4	40946	9	AC002984	AC002984 Human DNA
22	270.4	13.4	41369	9	CH19F24590	AD001527 Homo sapi
23	270.2	13.4	148624	2	AC084815	AC084815 Homo sapi
24	270.2	13.4	163026	2	AC103883	AC103883 Homo sapi
C 25	270.2	13.4	175051	9	AC010531	AC010531 Homo sapi
C 26	270.2	13.4	191635	2	AC137760	AC137760 Homo sapi
27	270.2	13.4	235286	9	AC136285	AC136285 Homo sapi
C 28	270	13.4	132933	9	AL137002	AL137002 Human DNA
C 29	269.2	13.4	155645	2	AC021153	AC021153 Homo sapi
30	269.2	13.4	177962	9	AC008763	AC008763 Homo sapi
31	268.8	13.4	182547	9	AP001201	AP001201 Homo sapi
32	268.8	13.4	208726	9	AP006287	AP006287 Homo sapi
C 33	268.4	13.3	122302	9	AC003982	AC003982 Homo sapi
C 34	267.8	13.3	193609	9	AC024580	AC024580 Homo sapi
C 35	267.4	13.3	196533	9	CNS000YVI	AL096870 Human chr
C 36	267.2	13.3	2836	6	AX832906	Sequence
C 37	267.2	13.3	2836	9	AK094129	AK094129 Homo sapi
38	267.2	13.3	126414	9	AC108084	AC108084 Homo sapi
C 39	267	13.3	39351	9	AC020945	AC020945 Homo sapi
C 40	267	13.3	217280	2	AC019337	AC019337 Homo sapi
C 41	266.8	13.3	163038	2	AC135542	AC135542 Pan trogl
C 42	266.6	13.3	145009	9	AC106742	AC106742 Homo sapi
C 43	266.6	13.3	152464	2	AC027421	AC027421 Homo sapi
C 44	266.6	13.3	157243	9	AL353622	AL353622 Human DNA
45	266.6	13.3	157986	2	AC025860	AC025860 Homo sapi

ALIGNMENTS

RESULT 1
AC002345
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

AC002345 Homo sapiens chromosome 17 clone CTB-20D5 map 17, 8 unordered pieces.
AC002345.2 GI:13027555
HTG: HTGS_PHASE1; HTGS_FULTOP; HTGS_CANCELLED.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 98829)
Birren,B., Linton,L., Nusbaum,C. and Landier,E.
Homo sapiens chromosome 17, clone CTB-20D5
Unpublished
2 (bases 1 to 98829)
Hawkins,T.L., Reeve,M.P., Christoffersen,A., Birren,B.W., Fasnath,K.H., Lander,E.S., McKernan,K., Munro,C., Richardson,P., Barna,N., Chang,A., Cooke,P., Daly,M.J., Devon,K., Dewar,K., Forrest,C., Gage,D., Geraigery,K., Guitau,G., Hagos,B., Huang,J., Jacotot,L., Lane,M., Lee,K., MacKenzie,J., Marquis,N., McDermott,J., Molla,M., Moloney,N., Morrow,J., Nachman,A., Naylor,J., O'Connor,T., Olotu,A., Peterson,K., Rollins,G., Spencer,J., Stilwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L., Traish,A., Wilmer,F., Zemtseva,I. and Zody,M.
Submitted (17-JUL-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 98829)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Glende,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karats,K., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lechoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,

Meneus, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, D., Zimmer, A. and Zody, M.

Direct Submission

Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Feb 21, 2001 this sequence version replaced gi:2262095.

All repeats were identified using RepeatMasker.

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WtBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L115

Center clone name: 20_D_5

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 30263: contig of 30263 bp in length

* 30264 30363: gap of 100 bp

* 30364 36119: contig of 5756 bp in length

* 36120 36219: gap of 100 bp

* 36220 40234: contig of 4015 bp in length

* 40235 40334: gap of 100 bp

* 40335 51410: contig of 11076 bp in length

* 51411 51510: gap of 100 bp

* 51511 56945: contig of 5435 bp in length

* 56946 57045: gap of 100 bp

* 57046 66865: contig of 9820 bp in length

* 66866 68965: gap of 100 bp

* 68966 70772: contig of 3807 bp in length

* 70773 70872: gap of 100 bp

* 70873 98829: contig of 27957 bp in length.

----- Location/Qualifiers

1. 98829

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone="CTB-20D5"

/clone_lib="CITB Human BAC"

FEATURES

source

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 2011; DB 2; Length 98829;

Mismatches 0; Indels 0; Gaps 0;

Matches 2011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CGGGCCGTTTGAAGAGCTACCCCGGACACAGACCCGACACAGATACCTCCAGC 60

CGGGCCGTTTGAAGAGCTACCCCGGACACAGACCCGACACAGATACCTCCAGC 85459

TCCTCTCAACCCACCCCTTTCAGGGTTGGAGAACTTGGAGCATAACTTGTCCATGA 120

TCCTCTCAACCCACCCCTTTCAGGGTTGGAGAACTTGGAGCATAACTTGTCCATGA 85519

GGAAATCTCCACCCAGAAATGGGTCCTTTCTGGCCCCCAGCCAGCTCCACATTAGAACA 180

Db	85520	GGAAATCTCCACCCAGAAATGGGTCCTTTCTGGCCCCCAGCCAGCTCCACATTAGAACA	85579
QY	181	TGACAAATAGAGGGGAAATGGAAATAAAGAGAGAAACGGTTTTCAGGAGCAGGGTT	240
Db	85580	TGACAAATAGAGGGGAAATGGAAATAAAGAGAGAAACGGTTTTCAGGAGCAGGGTT	85639
QY	241	TGGCTACAAAGTTGTGGATGTGGGTACCCATGCGCAAGTGTGAGGGAGGCTGGCGGGTG	300
Db	85640	TGGCTACAAAGTTGTGGATGTGGGTACCCATGCGCAAGTGTGAGGGAGGCTGGCGGGTG	85699
QY	301	TGGTGGCTCATGCTCTTAATCCAGCAGCTTTGGAGGCGCCAGGTGAGTAGATCACTTGA	360
Db	85700	TGGTGGCTCATGCTCTTAATCCAGCAGCTTTGGAGGCGCCAGGTGAGTAGATCACTTGA	85759
QY	361	GCGGGGAGTTTGACACAGCCTGGCCAAACATGGTGAACCCCACTCTGTACTTAAATAACA	420
Db	85760	GCGGGGAGTTTGACACAGCCTGGCCAAACATGGTGAACCCCACTCTGTACTTAAATAACA	85819
QY	421	AAAGTTAGCTGGGCGTGGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA	480
Db	85820	AAAGTTAGCTGGGCGTGGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA	85879
QY	481	TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGCCCGCTCTCAATAATAAAT	540
Db	85880	TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGCCCGCTCTCTCAATAATAAAT	85939
QY	541	ACAAAAATAGTTTGGATGTGGTGTGATGCTCCCTGTAGTCTTAGCTGAGGAGGCTGA	600
Db	85940	ACAAAAATAGTTTGGATGTGGTGTGATGCTCCCTGTAGTCTTAGCTGAGGAGGCTGA	85999
QY	601	GATGAAAGGATTTGCTTGGAGCTGGAGGTCAAGCTGCAAGTGTGAGCCGAGATGGCGCACT	660
Db	86000	GATGAAAGGATTTGCTTGGAGCTGGAGGTCAAGCTGCAAGTGTGAGCCGAGATGGCGCACT	86059
QY	661	GCATCTCAGCTGGGCAACAGAGTGAGCCCTGTCTCAGAAAAAATAAAAAAATAA	720
Db	86060	GCATCTCAGCTGGGCAACAGAGTGAGCCCTGTCTCAGAAAAAATAAAAAAATAA	86119
QY	721	GGAGAGGAGAGAGACTCAAGCAGCCCTCAGAGGCTGTCTGAGGCCCTGAGGCTGTCTG	780
Db	86120	GGAGAGGAGAGAGACTCAAGCAGCCCTCAGAGGCTGTCTGAGGCCCTGAGGCTGTCTG	86179
QY	781	CAGCATGTGGCCCGCAGGCCGCGGACTCTGTAAAGCACTGTCTGAGAGCACTCCCATCT	840
Db	86180	CAGCATGTGGCCCGCAGGCCGCGGACTCTGTAAAGCACTGTCTGAGAGCACTCCCATCT	86239
QY	841	TTCTCCAGATTTCTTAGACCTGTGCTTATACAGTCACTTTTATGTGTTTGGCCAAAT	900
Db	86240	TTCTCCAGATTTCTTAGACCTGTGCTTATACAGTCACTTTTATGTGTTTGGCCAAAT	86299
QY	901	TATTCAGCTCTGAAATTTCTCTGAGCTCCCTTTTACAAGCAGAGGTGAGCTAAGGCTGA	960
Db	86300	TATTCAGCTCTGAAATTTCTCTGAGCTCCCTTTTACAAGCAGAGGTGAGCTAAGGCTGA	86359
QY	961	GCTCAAGCATTTCAAAACCCCTACAGATCTGACGAATGTGTGCGCACGTCCTCCGAAATA	1020
Db	86360	GCTCAAGCATTTCAAAACCCCTACAGATCTGACGAATGTGTGCGCACGTCCTCCGAAATA	86419
QY	1021	TGAACACCTTTATGGGATGGGGCTGGGAGCAAGGCGGGAGAGCCATCTCTCA	1080
Db	86420	TGAACACCTTTATGGGATGGGGCTGGGAGCAAGGCGGGAGAGCCATCTCTCTCA	86479
QY	1081	GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCCGGCTCAATGTGAGTCCCT	1140
Db	86480	GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCCGGCTCAATGTGAGTCCCT	86539
QY	1141	GCTGCAACATCACTGGGCACTGGGTCCCTTCATTTCTCTCAAGAGGTGCTGTGAACC	1200
Db	86540	GCTGCAACATCACTGGGCACTGGGTCCCTTCATTTCTCTCAAGAGGTGCTGTGAACC	86599
QY	1201	CCAGCCTAGGAAAGGTAGATCCCTGGAGGAGGAGGTAATGTGTGTGGGAGAGCCT	1260

Db 86600 CCAAGCCTAGAAAAAGGTAGATCCCTGGAGGAGCAGGTAATGTGTTGGGAGAGCCT 86659
QY 1261 GGCCTGTGTCCTCTCTAGCTATGTAGATGACAGGGAGCTCGTGGAGTCTATGTACGAG 1320
Db 86660 GGCCTGTGTCCTCTCTAGCTATGTAGATGACAGGGAGCTCGTGGAGTCTATGTACGAG 86719
QY 1321 ACACCATCCCTGGAGCAAGACCTGGAGCGGCTCTTCCAGGAGCTGCAGGCACCTCTACCTC 1380
Db 86720 ACACCATCCCTGGAGCAAGACCTGGAGCGGCTCTTCCAGGAGCTGCAGGCACCTCTACCTC 86779
QY 1381 AACCTGATGCTTACGTCGCGCGGGGCTTGCACCGTCACTACGCGGCGCCAGCACATCAAC 1440
Db 86780 AACCTGATGCTTACGTCGCGCGGGGCTTGCACCGTCACTACGCGGCGCCAGCACATCAAC 86839
QY 1441 CTGAGGGGGCCCATCTCTGCTCACTGCTGGGTAAAGGGCACATGTCGGGCTTGAGGAGG 1500
Db 86840 CTGAGGGGGCCCATCTCTGCTCACTGCTGGGTAAAGGGCACATGTCGGGCTTGAGGAGG 86899
QY 1501 GTAAAGACGACACACAGTGTGAGTGGGGTTGGGACAGGGCTGACTAGAGGGTAGGGAGC 1560
Db 86900 GTAAAGACGACACACAGTGTGAGTGGGGTTGGGACAGGGCTGACTAGAGGGTAGGGAGC 86959
QY 1561 AGGCTGGGACTGAGAGACTCGAGCCCTGTGGGGGATGGTTGCCAGGCTGGAGGGGGGT 1620
Db 86960 AGGCTGGGACTGAGAGACTCGAGCCCTGTGGGGGATGGTTGCCAGGCTGGAGGGGGGT 87019
QY 1621 GGGCGCTGGAGTGGGAGGCCCCCACTGTCATCTGTCGCCACATTCATGCGAGATCTAT 1680
Db 87020 GGGCGCTGGAGTGGGAGGCCCCCACTGTCATCTGTCGCCACATTCATGCGAGATCTAT 87079
QY 1681 GTCCGCAAGTCAACATGATGGGGAAGGTAATATCTTGTCCAGGAGACCAACGCG 1740
Db 87080 GTCCGCAAGTCAACATGATGGGGAAGGTAATATCTTGTCCAGGAGACCAACGCG 87139
QY 1741 ACCATCAACATTTGTGTGATCTTAAAGGCGGAGGAGGCTGTGAGTGGGAGCTGGG 1800
Db 87140 ACCATCAACATTTGTGTGATCTTAAAGGCGGAGGAGGCTGTGAGTGGGAGCTGGG 87199
QY 1801 GAGGCTTTCCCAAGAGTGGCTGTGAGGAGGCGCTCGGAAGATGACAGGTTTGACAGA 1860
Db 87200 GAGGCTTTCCCAAGAGTGGCTGTGAGGAGGCGCTCGGAAGATGACAGGTTTGACAGA 87259
QY 1861 TGGAGTGGGGATGAGAGGACAGCGAGTGTTCAGGCCAAGGAACTGGAAACAAAGA 1920
Db 87260 TGGAGTGGGGATGAGAGGACAGCGAGTGTTCAGGCCAAGGAACTGGAAACAAAGA 87319
QY 1921 AGAAGCTGAGATGTAATCTACTTCAACCTGGAGCCCTCTTTGCAAGGCTGCAATC 1980
Db 87320 AGAAGCTGAGATGTAATCTACTTCAACCTGGAGCCCTCTTTGCAAGGCTGCAATC 87379
QY 1981 TCAGATGCCCTGAATGTGGAAGTAGGCGGT 2011
Db 87380 TCAGATGCCCTGAATGTGGAAGTAGGCGGT 87410

RESULT 2
AC113554/c
LOCUS
DEFINITION Homo sapiens chromosome 17, clone CTD-2501B8, complete sequence.
ACCESSION AC113554
VERSION AC113554.9 GI:24414535
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 179509)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone CTD-2501B8
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 179509)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Chopel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Farreira,P., Fitzhugh,W., Gage,D., Gallagan,J., Gardyna,S.,
Gard,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetty,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 179509)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Farreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Roh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 179509)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Farreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Roh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (26-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Oct 26, 2002 this sequence version replaced gi:22267826.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

[illegible]

```

116..324
/rpt_family="L1D2_Random_BK"
/rpt_family="MIR"
complement(716..1014)
/rpt_family="AluSx"
complement(1392..1690)
/rpt_family="AluSx"
complement(1715..2012)
/rpt_family="AluUb"
complement(2023..2322)
/rpt_family="AluSx"
complement(2483..2791)
/rpt_family="AluY"
2921..3229
/rpt_family="AluSx"
complement(3346..3454)
/rpt_family="MIR"
complement(3455..3774)
/rpt_family="AluSx"
complement(3775..3789)
/rpt_family="MIR"
complement(5114..5177)
/rpt_family="MIR"
5577..5875
/rpt_family="AluSg"
5876..5896
/rpt_family="(GA)n"
5928..6072
/rpt_family="MIR"
6083..6105
/rpt_family="AT_Rich"
6118..6267
/rpt_family="AluSq/x"
complement(6267..7050)
/rpt_family="L1MB8"
complement(7056..7198)
/rpt_family="FRAM"
complement(7199..7376)
/rpt_family="L1MB8"
7377..7675
/rpt_family="AluY"
complement(7688..8412)
/rpt_family="LTR65"
8644..8799
/rpt_family="MLTJ2"
complement(8828..9134)
/rpt_family="AluSx"
9135..9224
/rpt_family="MLTI1"
9362..9656
/rpt_family="AluSp"
10157..10472
/rpt_family="I12"
10593..10745
/rpt_family="L1PB2"
10752..11059
/rpt_family="AluSg"
11091..11115
```

QY	1	CCGCGCCGTTTGAAGAGCTCACCCCGGACACAAGAGGCCCGCACACAGATACCTCCACG	60
Db	118083		
		CCGCGCCGTTTGAAGAGCTCACCCCGGACACAAGAGGCCCGCACACAGATACCTCCACG	118024
QY	61	TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTGCTTCCATGA	120
Db	118023		
		TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTGCTTCCATGA	117964
QY	121	GGAACTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCACACATTAGAACAA	180
Db	117963		
		GGAACTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCACACATTAGAACAA	117904
QY	181	TGACAAATAGAAGGGGAAATGGAAAAATAACAGGAGAAACGGTTTTCCACGACACAGGTT	240
Db	117903		
		TGACAAATAGAAGGGGAAATGGAAAAATAACAGGAGAAACGGTTTTCCACGACACAGGTT	117844
QY	241	TGGCCTACAAGTTGTGATFGGGTACCCATGCCAAAGTGTGAGGGGAGGCTGGCCGGGTG	300
Db	117843		
		TGGCCTACAAGTTGTGATFGGGTACCCATGCCAAAGTGTGAGGGGAGGCTGGCCGGGTG	117784
QY	301	TGTFGGCTCATGCTCTTAATCCACAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTCAG	360
Db	117783		
		TGTFGGCTCATGCTCTTAATCCACAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTCAG	117724
QY	361	GCCGGGAGTTTGAGACCAAGCTGCCAACATGGTGAAACCCCATCTGTACTAAAAATACA	420
Db	117723		
		GCCGGGAGTTTGAGACCAAGCTGCCAACATGGTGAAACCCCATCTGTACTAAAAATACA	117664
QY	421	AAAGTTAGCTGGGGCTGGTGTAGATGCCTGTAGTCCAGACTACTTGGGAGGCTCAGGCA	480

Db 117663 AAAGTTAGCTGGCGCTGGTGTAGATGCCCTGTAGTCCAGCACTATTGGGAGGCTGAGGCA 117604
QY 481 TGAGAAATCGCTTCAGCCAGCCCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 540
Db 117603 TGAGAAATCGCTTCAGCCAGCCCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 117544
QY 541 ACAAAAATTTAGTTGGATGTGGTGTGGTGCATGCTGTAGTCCCTAGTCTGCTAGGAGGCTGA 600
Db 117543 ACAAAAATTTAGTTGGATGTGGTGTGGTGCATGCTGTAGTCCCTAGTCTGCTAGGAGGCTGA 117484
QY 601 GATGGAGAGATTCCTTGAGCTGGGAGGCTCAAGGCTGCAGTGAGCCGAGATGGCGCACT 660
Db 117483 GATGSAAGGATTCCTTGAGCTGGGAGGCTCAAGGCTGCAGTGAGCCGAGATGGCGCACT 117424
QY 661 GCACCTCAGGCTGGGCAACAGAGTGAGAGCCCTCTCTCAGAAAAAANAANAANAANA 720
Db 117423 GCACCTCAGGCTGGGCAACAGAGTGAGAGCCCTCTCTCAGAAAAAANAANAANAANA 117364
QY 721 GGAGGAGAGAGACTCAAGCACGCCCTCTCAGGAGTCTGCTGAGGCCCTGCGAGGTGTCTG 780
Db 117363 GGAGGAGAGAGACTCAAGCACGCCCTCTCAGGAGTCTGCTGAGGCCCTGCGAGGTGTCTG 117304
QY 781 CAGCATGTGCCCCCAGCGCGGAGCTCTGTAAGCCACTGCTGAGAGGCACTCCCATCT 840
Db 117303 CAGCATGTGCCCCCAGCGCGGAGCTCTGTAAGCCACTGCTGAGAGGCACTCCCATCT 117244
QY 841 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGTTTGCCTAAATTT 900
Db 117243 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGTTTGCCTAAATTT 117184
QY 901 TATTCAGCTCTGAATTTCTCTGAGCTCCCTTTACAAGCAGAGGTGAGTAAAGGCTGGA 960
Db 117183 TATTCAGCTCTGAATTTCTCTGAGCTCCCTTTACAAGCAGAGGTGAGTAAAGGCTGGA 117124
QY 961 GCTCAAGGATTCAAACCCCTACAGATCTGAGAAATGTATGGGCCACGTCCTCGGAAATA 1020
Db 117123 GCTCAAGGATTCAAACCCCTACAGATCTGAGAAATGTATGGGCCACGTCCTCGGAAATA 117064
QY 1021 TGAAGACCTGTTATGGCATGGGAGGCTGGGAGACAGGCGGGGAGAGCCATCTCCA 1080
Db 117063 TGAAGACCTGTTATGGCATGGGAGGCTGGGAGACAGGCGGGGAGAGCCATCTCCA 117004
QY 1081 GTTTTACCGGAAATAGTGGAACTCATCAACAGGCTGCCCGCTCAATGGTGGTCCCT 1140
Db 117003 GTTTTACCGGAAATAGTGGAACTCATCAACAGGCTGCCCGCTCAATGGTGGTCCCT 116944
QY 1141 GCTGCCAATCACTGGGCACTGGGTCCTTCAATTTCTCAAGAGGTGCTGTGAAC 1200
Db 116943 GCTGCCAATCACTGGGCACTGGGTCCTTCAATTTCTCAAGAGGTGCTGTGAAC 116884
QY 1201 CCAAGCCTAGGAAAGGTAGATCCCTGGAGGAGGAGGTAATGTGGTGGGAGGCT 1260
Db 116883 CCAAGCCTAGGAAAGGTAGATCCCTGGAGGAGGAGGTAATGTGGTGGGAGGCT 116824
QY 1261 GCGTGTGCTCCCTCTGTAGGCTATGTAGATGCAAGGGAATCTGGAGGCTATGTACGAG 1320
Db 116823 GCGTGTGCTCCCTCTGTAGGCTATGTAGATGCAAGGGAATCTGGAGGCTATGTACGAG 116764
QY 1321 ACACCATCCTGGAGCAACCTGGAGGCTCTTCCAGAGTGCAGGCACTTACCTC 1380
Db 116763 ACACCATCCTGGAGCAACCTGGAGGCTCTTCCAGAGTGCAGGCACTTACCTC 116704
QY 1381 AACCTCATGCTTACCTGCGCGGCCCTCGACCCGTCATCAGGGCCCGAGCACATCAAC 1440
Db 116703 AACCTCATGCTTACCTGCGCGGCCCTCGACCCGTCATCAGGGCCCGAGCACATCAAC 116644
QY 1441 CTGGAGGGGCCCATTCCTGCTCACTGCTGGGTAAAGGCACATGTGGGCTTTAGGAGG 1500
Db 116643 CTGGAGGGGCCCATTCCTGCTCACTGCTGGGTAAAGGCACATGTGGGCTTTAGGAGG 116584
QY 1501 GTAAGAAGGACACAGTGTAGTGGGTTGGGACAGGGCTGACTAGAGGGTAGGGAGC 1560
Db 116583 GTAAGAAGGACACAGTGTAGTGGGTTGGGACAGGGCTGACTAGAGGGTAGGGAGC 116524

QY 1561 AGGCTGGGACTGAGAGACTCCAGCCCTGTGGGGATGTTGCCAGGCTGGAGGGGGT 1620
Db 116523 AGGCTGGGACTGAGAGACTCCAGCCCTGTGGGGATGTTGCCAGGCTGGAGGGGGT 116464
QY 1621 GGGCGCTGGAGTGGGGAGCCCCCCTTGCATCTGGTGCCACATTCATCTGCAGATCTAT 1680
Db 116463 GGGCGCTGGAGTGGGGAGCCCCCCTTGCATCTGGTGCCACATTCATCTGCAGATCTAT 116404
QY 1681 GTGCGGCAAGTCAACATGATGGGGAGGAAGTTAATAATCTTTGTCCAGAGACCAAGGC 1740
Db 116403 GTGCGGCAAGTCAACATGATGGGGAGGAAGTTAATAATCTTTGTCCAGAGACCAAGGC 116344
QY 1741 ACCCATCAACATCTGTGTATCTTAGAGGGCGAGAGAGGCTGTAGTGGGAGCTGG 1800
Db 116343 ACCCATCAACATCTGTGTATCTTAGAGGGCGAGAGAGGCTGTAGTGGGAGCTGG 116284
QY 1801 GAGGCTTTTCCCAAGAGGTGGCTGTGAGCAGAGGCTTCGGAAGATGACAGGTTTGACAGA 1860
Db 116283 GAGGCTTTTCCCAAGAGGTGGCTGTGAGCAGAGGCTTCGGAAGATGACAGGTTTGACAGA 116224
QY 1861 TGGGAAGTGGGGATGAGAGGACAGACGAGTGTTCAGGCCAAGGAACTGGGAACAAAGA 1920
Db 116223 TGGGAAGTGGGGATGAGAGGACAGACGAGTGTTCAGGCCAAGGAACTGGGAACAAAGA 116164
QY 1921 AGAACCTGAGATGTAAATCTACTTCAACCTGGACCTCTTGCAGAGGCTGCAATC 1980
Db 116163 AGAACCTGAGATGTAAATCTACTTCAACCTGGACCTCTTGCAGAGGCTGCAATC 116104
QY 1981 TCAGATGCCCTGAATGTGTGAAGTAGGCGGT 2011
Db 116103 TCAGATGCCCTGAATGTGTGAAGTAGGCGGT 116073

RESULT 3
AY436326
LOCUS AY436326 23424 bp DNA linear PRI 26-OCT-2003
DEFINITION Homo sapiens angiotensin I converting enzyme (peptidyl-dipeptidase
A) 1 (ACE) gene, complete cds.
ACCESSION AY436326
VERSION AY436326.1 GI:37790803
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 23424)
AUTHORS Rieder M.J., da Ponte, S.H., Kuldane, S.A., Rajkumar, N., Smith, J.D.,
Toth, E.J., Krauss, R.M. and Nickerson, D.A.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-2003) Genome Sciences, University of Washington,
1705 NE Pacific, Seattle, WA 98195, USA
FEATURES
Location/Qualifiers
1..23424
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
54..205
/rpt_family="12"
/rpt_type="dispersed"
1024..1143
/note="Region not scanned for variation"
1502..1593
/rpt_family="MIR"
/rpt_type="dispersed"
1688..1942
/note="Region not scanned for variation"
2097..2357
/note="Region not scanned for variation"
/note="Region not scanned for variation"
/gene="ACE"
2355..22724
mRNA
5618..5809,6372..6469,6847..7019,7747..7970,8310..8454,

variation	/frequency="0.17"				8581			
	/replaces="t"				/gene="ACE"			
	/frequency="0.03"				/replaces="t"			
repeat_region	8608. .8650				/rpt_family="L2"			
variation	/rpt_type=dispersed				8683			
	/gene="ACE"				/frequency="0.43"			
variation	/replaces="t"				8824			
	/gene="ACE"				/frequency="0.01"			
	/replace="c"							
Query Match 99.3%; Score 1997.4; DB 9; Length 23424;								
Best Local Similarity 99.9%; Pred. No. 0;								
Matches 2009; Conservative 0; Mismatches 1; Indels 1; Gaps 1;								
QY	1	CGGCGCGTTTGAAGAGCTCACCCCGGACACAAGGACCCGGACACAGATACCTCCAGC	60					
DB	12944	CGAGCCCGTTTGAAGAGCTCACCCCGGACACAAGGACCCGGACACAGATACCTCCAGC	13003					
QY	61	TCCTCTCAACCCACCCTTTCAGGGTTGGAGAACTTGAGGCATAAACTTGTCTCCATGA	120					
DB	13004	TCCTCTCAACCCACCCTTTCAGGGTTGGAGAACTTGAGGCATAAACTTGTCTCCATGA	13063					
QY	121	GGAACTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCACTTTAGAACAA	180					
DB	13064	GGAACTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCACTTTAGAACAA	13123					
QY	181	TGACAAATAGAGGGGAAATGGAAAAATAACAGGAGAAACGGTTTTCCAGGACAGGGTT	240					
DB	13124	TGACAAATAGAGGGGAAATGGAAAAATAACAGGAGAAACGGTTTTCCAGGACAGGGTT	13183					
QY	241	TGGCCTTACAAGTTGTGGATGTGGGTACCCATGCGAAGTGTGAGGGAGGCTGGCCGGGTG	300					
DB	13184	TGGCCTTACAAGTTGTGGATGTGGGTACCCATGCGAAGTGTGAGGGAGGCTGGCCGGGTG	13243					
QY	301	TGTTGGCTCATGCTCTAATCCAGCACATTGGGAGGCCAAGGTGATGATCACTTTGAG	360					
DB	13244	TGTTGGCTCATGCTCTAATCCAGCACATTGGGAGGCCAAGGTGATGATCACTTTGAG	13303					
QY	361	GCGGGAGTTTGAGACAGCTGGGCCAACAATGGTGAAACCCCATCTGTACTAAAAATACA	420					
DB	13304	GCGGGAGTTTGAGACAGCTGGGCCAACAATGGTGAAACCCCATCTGTACTAAAAATACA	13363					
QY	421	AAAGTTAGCTGGGCGTGGTGTAGATGCCCTGTAGTCCCAAGCTACTTGGGAGGCTGAGGCA	480					
DB	13364	AAAGTTAGCTGGGCGTGGTGTAGATGCCCTGTAGTCCCAAGCTACTTGGGAGGCTGAGGCA	13423					
QY	481	TGAGATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT	540					
DB	13424	TGAGATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT	13483					
QY	541	ACAAAAATTTAGTTGATGTGGTGGTGCATGCTGTAGTCTTAGCTGCTAGGAGGCTGA	600					
DB	13484	ACAAAAATTTAGTTGATGTGGTGGTGCATGCTGTAGTCTTAGCTGCTAGGAGGCTGA	13543					
QY	601	GATGGAAGGATTCCTTGAGCCTGGGAGGTCAGGCTGCAGTGCAGATGGCGCCACT	660					
DB	13544	GATGGAAGGATTCCTTGAGCCTGGGAGGTCAGGCTGCAGTGCAGATGGCGCCACT	13603					
QY	661	GCACTCCAGCTGGGCAACAGAGTGAAGCCCTGCTCAGAAAAAATAAAAAAATAA	720					
DB	13604	GCACTCCAGCTGGGCAACAGAGTGAAGCCCTGCTCAGAAAAAATAAAAAAATAA	13663					
QY	721	GGAGAGAGAGACTCAAGCACGCCCTTCACAGGACTGCTGAGGCCCTGCAGGTGTCTG	780					
DB	13664	GGAGAGAGAGACTCAAGCACGCCCTTCACAGGACTGCTGAGGCCCTGCAGGTGTCTG	13723					
QY	781	CAGCATGTGCCCCCAGGCGGGGACTCTGTAAAGCCACTGCTGGAGAGCCACTCCCACCT	840					
DB	13724	CAGCATGTGCCCCCAGGCGGGGACTCTGTAAAGCCACTGCTGGAGAGCCACTCCCACCT	13783					
QY	841	TTCTCCCATTTTCTCTAGACCTGCTGCTATACAGTCACTTTTATGTGGTTTCGCCAATTT	900					
DB	13784	TTCTCCCATTTTCTCTAGACCTGCTGCTATACAGTCACTTTTATGTGGTTTCGCCAATTT	13843					
QY	901	TATTCAGCTCTGAAATTTCTCTGAGCTCCCTTACAGCAGAGGTGAGCTAAGGGCTGGA	960					
DB	13844	TATTCAGCTCTGAAATTTCTCTGAGCTCCCTTACAGCAGAGGTGAGCTAAGGGCTGGA	13903					
QY	961	GCTCAAGGCATTCAAAACCCCTACAGATCTGACGAAATGTGATGGCCACGTCCTCCGGAATA	1020					
DB	13904	GCTCAAGGCATTCAAAACCCCTACAGATCTGACGAAATGTGATGGCCACGTCCTCCGGAATA	13963					
QY	1021	TGAAGACCTGTTATGGGCATGGGAGGCTGGGAGACAAGGGGGGAGAGCCATCCTCCA	1080					
DB	13964	TGAAGACCTGTTATGGGCATGGGAGGCTGGGAGACAAGGGGGGAGAGCCATCCTCCA	14023					
QY	1081	GTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCCGCTCAATGTGTAGTCCCT	1140					
DB	14024	GTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCCGCTCAATGTGTAGTCCCT	14082					
QY	1141	GCTGCCAACATCACTGGCACTTGGGTCCCTTCAATTTCTCAAGAGGTGCTGTGAAACC	1200					
DB	14083	GCTGCCAACATCACTGGCACTTGGGTCCCTTCAATTTCTCAAGAGGTGCTGTGAAACC	14142					
QY	1201	CCAAGCTTAGGAAAAGGTAGATCCCTGGAGAGGAGCAGGTAAATGTGTGTGGGAGAGCCT	1260					
DB	14143	CCAAGCTTAGGAAAAGGTAGATCCCTGGAGAGGAGCAGGTAAATGTGTGTGGGAGAGCCT	14202					
QY	1261	GGCTGTGTCCCTCTGTAGGCTATGTAGATGACAGGGGACTCGTGGAGGCTCTATGTACGAG	1320					
DB	14203	GGCTGTGTCCCTCTGTAGGCTATGTAGATGACAGGGGACTCGTGGAGGCTCTATGTACGAG	14262					
QY	1321	ACACCATCCCTGGAGCAACCTGGAGCGGCTTTCACAGAGCTGCAGCACCTCTACCTC	1380					
DB	14263	ACACCATCCCTGGAGCAACCTGGAGCGGCTTTCACAGAGCTGCAGCACCTCTACCTC	14322					
QY	1381	AACTGTCATGCCCTACGTGGCGCGGCGCTGCACCGTCACTACGGGGCCAGCACATCAAC	1440					
DB	14323	AACTGTCATGCCCTACGTGGCGCGGCGCTGCACCGTCACTACGGGGCCAGCACATCAAC	14382					
QY	1441	CTGGAGGGGCCAATTCCTGCTCACTGCTGGTGAAGGGCACAATGTCGGGCTTGAGGAGG	1500					
DB	14383	CTGGAGGGGCCAATTCCTGCTCACTGCTGGTGAAGGGCACAATGTCGGGCTTGAGGAGG	14442					
QY	1501	GTAAAGACGGACCCACAGTGTGAGTGGGTTGGGACAGGCTGACTAGAGGGTAGGGAGC	1560					
DB	14443	GTAAAGACGGACCCACAGTGTGAGTGGGTTGGGACAGGCTGACTAGAGGGTAGGGAGC	14502					
QY	1561	AGGCTGGGAGCTGAGAGACTCCAGGCCCTGTGGGGATGTGTCACAGGCTGAGAGGGGGT	1620					
DB	14503	AGGCTGGGAGCTGAGAGACTCCAGGCCCTGTGGGGATGTGTCACAGGCTGAGAGGGGGT	14562					
QY	1621	GGGCGCTGGGAGTGGGAGCCCCCACTTGTGATCTGTGTGCCACATTCATGACAGATCTAT	1680					
DB	14563	GGGCGCTGGGAGTGGGAGCCCCCACTTGTGATCTGTGTGCCACATTCATGACAGATCTAT	14622					
QY	1681	GTGCGGCAAGTCAACATGGATGGGGGAGAGGTTAATAATCTTGTCCAGGAGCACCGGC	1740					
DB	14623	GTGCGGCAAGTCAACATGGATGGGGGAGAGGTTAATAATCTTGTCCAGGAGCACCGGC	14682					
QY	1741	ACCCATCAACAATTTGTGTGATCTTAGAGGGCAGGAAGGCTGTGAGTGGGAGCTGGG	1800					
DB	14683	ACCCATCAACAATTTGTGTGATCTTAGAGGGCAGGAAGGCTGTGAGTGGGAGCTGGG	14742					
QY	1801	GAGGCTTTGCCAAGAGGTGGCTGTGAGCAGGGCTCGGAAGATGACAGGGTTTGACAGA	1860					
DB	14743	GAGGCTTTGCCAAGAGGTGGCTGTGAGCAGGGCTCGGAAGATGACAGGGTTTGACAGA	14802					

variation	5492	/gene="DCP1"	QY	1	CCGGCCCGTTTGAAGAGCTCACCCCGACACAGGACCGGCACAGATACCTCCAGC	60
		/frequency="0.09"				
variation	5878	/replace="t"	Db	13228	CCAGCCCGTTTGAAGAGCTCACCCCGACACAGGACCGGCACAGATACCTCCAGC	13287
		/gene="DCP1"	QY	61	TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTGTCTCCATGA	120
		/frequency="0.09"				
variation	6028	/replace="c"	Db	13288	TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTGTCTCCATGA	13347
		/gene="DCP1"	QY	121	GGAACTCTCACCCAGAAATGGGTCTTTCTGGGCCCCAGCCAGCTCCCACTTTAGAACAA	180
		/frequency="0.05"				
variation	6029	/replace="t"	Db	13348	GGAACTCTCACCCAGAAATGGGTCTTTCTGGGCCCCAGCCAGCTCCCACTTTAGAACAA	13407
		/gene="DCP1"	QY	181	TGACAAATAGAGGGGAATGGAATAAACAGGAGAAAGGTTTCCAGGACAGGGTT	240
		/frequency="0.05"				
variation	6145	/replace="t"	Db	13408	TGACAAATAGAGGGGAATGGAATAAACAGGAGAAAGGTTTCCAGGACAGGGTT	13467
		/gene="DCP1"	QY	241	TGGCCTTACAAAGTTGTGGATGTGGGTACCCATGCTCAAGTGTGAGGGAGGCTGGCCGGGTG	300
		/frequency="0.05"				
repeat_region		/replace="t"	Db	13468	TGGCCTTACAAAGTTGTGGATGTGGGTACCCATGCTCAAGTGTGAGGGAGGCTGGCCGGGTG	13527
		complement (<6245. .>6320)				
		/note="putative"	QY	301	TGGTGGCTCATGCTCTAATCCAGACCTTTGGGAGGCCAAGGTGAGTAGATCACTTGAG	360
		/rpt_family="MIR"				
variation	6435	/rpt_type=dispersed	Db	13528	TGGTGGCTCATGCTCTAATCCAGACCTTTGGGAGGCCAAGGTGAGTAGATCACTTGAG	13587
		/gene="DCP1"	QY	361	GCCGGGAGTTTGAGACCAGCCTGGCCCAACATGGTGAACCCCATCTGTACTAAAAATACA	420
		/frequency="0.32"				
variation	6604	/replace="g"	Db	13588	GCCGGGAGTTTGAGACCAGCCTGGCCCAACATGGTGAACCCCATCTGTACTAAAAATACA	13647
		/gene="DCP1"	QY	421	AAAGTTAGCTGGGCGTGGTGGTAGATGCCTGTAGTCCCACTTGGGAGGCTGAGGCA	480
		/frequency="0.05"				
repeat_region		/replace="a"	Db	13648	AAAGTTAGCTGGGCGTGGTGGTAGATGCCTGTAGTCCCACTTGGGAGGCTGAGGCA	13707
		<7584. .>7889				
		/note="putative"	QY	481	TGAGAAATCGCTTGAGCCCAAGCCTGGGCAATAACAGCAAGACCCGCTCTTACAAATAAAT	540
		/rpt_family="Alu"				
variation	7693	/rpt_type=dispersed	Db	13708	TGAGAAATCGCTTGAGCCCAAGCCTGGGCAATAACAGCAAGACCCGCTCTTACAAATAAAT	13767
		/gene="DCP1"	QY	541	ACAAAAAATTAGTTGGATGTGGTGGTGCATGCTGTAGTCCCTAGTCTGCTAGGAGGCTGA	600
		/frequency="0.05"				
variation	7831	/replace="a"	Db	13768	ACAAAAAATTAGTTGGATGTGGTGGTGCATGCTGTAGTCCCTAGTCTGCTAGGAGGCTGA	13827
		/gene="DCP1"	QY	601	GATGGAAGGATTGCTTGAGCCTGGAGGTCAGGCTGACAGTGCAGCGAGATGGCGCACT	660
		/frequency="0.23"				
variation	8128	/replace="a"	Db	13828	GATGGAAGGATTGCTTGAGCCTGGAGGTCAGGCTGCAAGTGCAGCGAGATGGCGCACT	13887
		/gene="DCP1"	QY	661	GCACTCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTCAG- -AAAGAAAAAATAAAAAA	720
		/frequency="0.18"				
repeat_region		/replace="t"	Db	13888	GCACTCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTCAG- -AAAGAAAAAATAAAAAA	13945
		<8893. .>8941				
		/note="putative"	QY	721	GGAGAGGAGAGAGACTCAAGCACGCCCTCAGAGGACTGCTGAGGCCCTGCGAGGTCTCG	780
		/rpt_family="L2"				
variation	8935	/rpt_type=dispersed	Db	13946	GGAGAGGAGAGAGACTCAAGCACGCCCTCAGAGGACTGCTGAGGCCCTGCGAGGTCTCG	14005
		/gene="DCP1"	QY	781	CAGCATGTGGCCCGGCGGGAGCTCTGTAAGCCACTGCTGGAGCCACTCCCATCTCT	840
		/frequency="0.05"				
variation	8968	/replace="c"	Db	14006	CAGCATGT- -GCCCGAGGCGGGGAGCTCTGTAAGCCACTGCTGGAGA- -CCACTCCCATCT	14063
		/gene="DCP1"	QY	841	TTCTCCCATTTCTCTAGACCTGCTGCCT- -	868
		/frequency="0.36"				
variation	9191	/replace="t"	Db	14064	TTCTCCCATTTCTCTAGACCTGCTGCCTATACAGTCACTTTTTTTTTTTTTTTTGAGACGG	14123
		/gene="DCP1"	QY	869	-----	868
		/frequency="0.09"				
variation	9222	/replace="t"	Db	14124	AGTCTCGTCTGTGTCGCCAGGCTGGAGTGCAGTGGCGGAGTCTCGGCTCACTGCAAGCTC	14183
		/gene="DCP1"	QY	869	-----	868
		/frequency="0.09"				
variation	9509	/replace="t"	Db	14184	CGCTCCCGGGTTACGGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACACAGGC	14243
		/gene="DCP1"	QY	869	-----	868
		/frequency="0.09"				
variation	9509	/replace="t"	Db	14244	GCCCCCAGCTACGCCCGGCTAAATTTTTTTTGTATTTTTTTAGTAGACGGGGTTTTCACCGTTT	14303

Query Match 82.9%; Score 1667.6; DB 9; Length 24070;
 Best Local Similarity 87.1%; Pred. No. 0;
 Matches 2003; Conservative 4; Indels 293; Gaps 4;

```
QY 869 ----- 868
Db 14304 TAGCGGGATGGTCTCTGATCTCTGACCTCGTATCCGCGGCTCCGCGCTCCCAAAGTG 14363
QY 869 -----ATACAGTACATTTTATGTGGTTTCGCCAATTTTATTCAGCTC 911
Db 14364 CTGGGATTTACAGGCGTGATACAGTACATTTTATGTGGTTTCGCCAATTTTATTCAGCTC 14423
QY 912 TGAATTTCTTGAGTCCCTTTACAGCAGAGGTGAGCTAAGGGCTGGAGCTCAAGGCAT 971
Db 14424 TGAATTTCTTGAGTCCCTTTACAGCAGAGGTGAGCTAAGGGCTGGAGCTCAAGGCAT 14483
QY 972 TCAAAACCCCTACAGATCTGAGGAATGTGATGGCCACGCTCCCGGAAATATGAAGCCTGT 1031
Db 14484 TCAAAACCCCTACAGATCTGAGGAATGTGATGGCCACGCTCCCGGAAATATGAAGCCTGT 14543
QY 1032 TATGGGCATGGAGGCTGGGAGACAAAGGGGGAGAGCCATCTCCAGTTTACCCGA 1091
Db 14544 TATGGGCATGGAGGCTGGGAGACAAAGGGGGAGAGCCATCTCCAGTTTACCCGA 14603
QY 1092 AATAGTGGAACTCATCAACAGGCTGCCGGCTCAATGGTGAGTCCCTGCTGCCAAT 1151
Db 14604 AATAGTGGAACTCATCAACAGGCTGCCGGCTCAATGGTGAGTCCCTGCTGCCAAT 14663
QY 1152 CACTGGCACTGGGTCCTTCAATTTCTCAAGAGGTGCTGTGAAACCCCAAGCCTAGG 1211
Db 14664 CACTGGCACTGGGTCCTTCAATTTCTCAAGAGGTGCTGTGAAACCCCAAGCCTAGG 14723
QY 1212 AAAAGTGTAGTCCCTGGAGGAGGAGTAATCTGTGTGTGGAGAGCCTGGCTGTGCC 1271
Db 14724 AAAAGTGTAGTCCCTGGAGGAGGAGTAATCTGTGTGTGGAGAGCCTGGCTGTGCC 14783
QY 1272 CTCTGTAGGCTATGTAGATGCAAGGGACTCGTGGAGGTCTATGTACGAGACACCATCCT 1331
Db 14784 CTCTGTAGGCTATGTAGATGCAAGGGACTCGTGGAGGTCTATGTACGAGACACCATCCT 14843
QY 1332 GGAGCAAGACTGGAGCGGCTCTTCAGAGCTGAGGACCTCTACCTCAACCTGCATGC 1391
Db 14844 GGAGCAAGACTGGAGCGGCTCTTCAGAGCTGAGGACCTCTACCTCAACCTGCATGC 14903
QY 1392 CTACGTGGCGGGCCCTGCACCTGCTACCTACGGGGCCAGACATCAACCTGGAGGGCC 1451
Db 14904 CTACGTGGCGGGCCCTGCACCTGCTACCTACGGGGCCAGACATCAACCTGGAGGGCC 14963
QY 1452 CATTCCTGCTCACTGCTGGTAAAGGCACATGTCTGGGCTTTGAGAGGGTTAAAGACGA 1511
Db 14964 CATTCCTGCTCACTGCTGGTAAAGGCACATGTCTGGGCTTTGAGAGGGTTAAAGACGA 15023
QY 1512 CCACAGTGTAGTGGGTTGGGACAGGCTGACTAGAGGTAGGAGCAGGCTGGGGAC 1571
Db 15024 CCACAGTGTAGTGGGTTGGGACAGGCTGACTAGAGGTAGGAGCAGGCTGGGGAC 15083
QY 1572 TCAGAGACTCCAGCCCTGTGGGGGATGTTGCCAGGCTGGAGGGGGTGGCGCTGGGA 1631
Db 15084 TCAGAGACTCCAGCCCTGTGGGGGATGTTGCCAGGCTGGAGGGGGTGGCGCTGGGA 15143
QY 1632 GTGGGAGCCCCCACTTGTCATCTGGTGCCACATTCACCTGCAGATCTATGTGGGCAAGT 1691
Db 15144 GTGGGAGCCCCCACTTGTCATCTGGTGCCACATTCACCTGCAGATCTATGTGGGCAAGT 15203
QY 1692 CACCATGGATGGGGAGAGTAATATCTTGTCCAGAGACCCAGGACCCATCACAA 1751
Db 15204 CACCATGGATGGGGAGAGTAATATCTTGTCCAGAGACCCAGGACCCATCACAA 15263
QY 1752 CATTTGTGATCTTACAGGGCCAGGAAGGCTGTGAGTGGGAGCTGGGAGGCTTTGCC 1811
Db 15264 CATTTGTGATCTTACAGGGCCAGGAAGGCTGTGAGTGGGAGCTGGGAGGCTTTGCC 15323
QY 1812 AAGAGTGGCCTGTGAGCAGGGCCTCGGAAGATGACAGGGTTTGACAGATGGGAAGTGG 1871
Db 15324 AAGAGTGGCCTGTGAGCAGGGCCTCGGAAGATGACAGGGTTTGACAGATGGGAAGTGG 15383
QY 1872 GGATGAGAGACAGACGCGAGTGTTCAGGCCAAGGGNACTGGACAAAGAACCTGAGA 1931
```

```
Db 15384 GGATGAGAGACAGACGCGAGTGTTCAGGCCAAGGGNACTGGACAAAGAACCTGAGA 15443
QY 1932 ATGTAATCTACTTCAACCCCTGGACCCCTCTCTTTGCAAGGGCTGCAATCTCAGATSCCT 1991
Db 15444 ATGTAATCTACTTCAACCCCTGGACCCCTCTCTTTGCAAGGGCTGCAATCTCAGATSCCT 15503
QY 1992 GAATGTGTGAAGTAGGCGGT 2011
Db 15504 GAATGTGTGAAGTAGGCGGT 15523

RESULT 5
A28005 Human ACE gene intron 15-18 sequence. linear PAT 18-SEP-1995
ACCESSION A28005
VERSION A28005.1 GI:1247488
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2649)
AUTHORS .
TITLE METHODS AND MEANS FOR STUDYING GENETIC POLYMORPHISM IN THE
ANGIOTENSIN CONVERTING ENZYME
JOURNAL Patent: WO 9300360-A 1 07-JAN-1993;
FEATURES
source
1. 2649
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 61.3%; Score 1233.6; DB 6; Length 2649;
Best Local Similarity 84.0%; Pred. No. 0; Mismatches 295; Gaps 6;
Matches 1594; Conservative 0; Indels 295; Gaps 6;
QY 1 CGGGCCCTTTGAAAGAGCTCACCCCGACACAAAGAGCCGACACAGATACCTCCAGC 60
Db 759 CGAGCCCTTTGAAAGAGCTCACCCCGACACAAAGAGCCGACACAGATACCTCCAGC 818
QY 61 TCCCTCTCAACCCACCTTTCCAGGGTTGGAGAACTTGAGGCATAACTTGTCTCATGA 120
Db 819 TCCCTCTCAACCCACCTTTCCAGGGTTGGAGAACTTGAGGCATAACTTGTCTCATGA 878
QY 121 GGAATCTCCACCCAGAAATGGTCTTTCTGGCCCCCAGCCAGCTCCACATTAGAACAA 180
Db 879 GGAATCTCCACCCAGAAATGGTCTTTCTGGCCCCCAGCCAGCTCCACATTAGAACAA 938
QY 181 TGACAAATAGAGGGGAAATGAAATAAACAAGAGAAACCGTTTTCAGGACAGGGTT 240
Db 939 TGACAAATAGAGGGGAAATGAAATAAACAAGAGAAACCGTTTTCAGGACAGGGTT 998
QY 241 TGGCCCTCAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCGGGTG 300
Db 999 TGGCCCTCAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCGGGTG 1058
QY 301 TGGTGGCTCATGCTTAATCCAGCACTTTTGGAGGGCCAGAGGTGAGTAGATCACTTGA 360
Db 1059 TGGTGGCTCATG-CTCTTAATCCAGCACTTTTGGAGGGCCAGAGGTGAGTAGATCACTTGA 1117
QY 361 GCCGGAGTTTGGAGCCAGCTGGCCAAATGATGGTGAACCCCATCTGTACTAAAAATACA 420
Db 1118 GCCGGAGTTTGGAGCCAGCTGGCCAAATGATGGTGAACCCCATCTGTACTAAAAATACA 1177
QY 421 AAGTTAGCTGGCGCTGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 480
Db 1178 AAGTTAGCTGGCGCTGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 1237
QY 481 TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCGCTCTTACAAATAAAT 540
```


Db 948 GCCGGAGTTTGGAGACCCAGCTGGCCAACTAGTGTAACCCCACTGTACTATAAAATACA 1007
 Qy 421 AAAGTTACTGGCGTGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGAGCTGAGGCA 480
 Db 1008 AAAGTTACTGGCGTGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGAGCTGAGGCA 1067
 Qy 481 TGAGAAATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT 540
 Db 1068 TGAGAAATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT 1127
 Qy 541 ACAAAAATTTAGTTGGATGTGGTGTGATGCTGTAGTCCAGCTACTTGGAGAGCTGGA 600
 Db 1128 ACAAAAATTTAGTTGGATGTGGTGTGATGCTGTAGTCCAGCTACTTGGAGAGCTGGA 1187
 Qy 601 GATGAAGGATTTGCTTGGAGCTGGGAGTCAAGGCTGAGGCTGAGGAGAGTGGCGCCACT 660
 Db 1188 GATGAAGGATTTGCTTGGAGCTGGGAGTCAAGGCTGAGGCTGAGGAGAGTGGCGCCACT 1247
 Qy 661 GCATCCAGCTTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAATAAAAAA 720
 Db 1248 GCATCCAGCTTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAATAAAAAA 1305
 Qy 721 GGAGGAGAGAGACTCAAGCAAGCCCTCACAGGACTGCTGAGGAGTGGCTG 780
 Db 1306 GGAGGAGAGAGACTCAAGCAAGCCCTCACAGGACTGCTGAGGAGTGGCTG 1365
 Qy 781 CAGCATGTGGCCCGCCAGGCGGGGACTCTGTAGCCACTGCTGGAGAGCACTCCCATCT 840
 Db 1366 CAGCATGTG--CCAGGCGGGGACTCTGTAGCCACTGCTGGAGAGCACTCCCATCT 1422
 Qy 841 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 891
 Db 1423 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 1473

RESULT 7
 AR137381
 LOCUS AR137381 1856 bp DNA linear PAT 16-JUN-2001
 DEFINITION Sequence 128 from patent US 6197505.
 ACCESSION AR137381
 VERSION AR137381.1 GI:14478890
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1856)
 AUTHORS Norberg,L.Torbjorn., Andersson,M.Kristina. and Lindstrom,P.Harry.Rutger.
 TITLE Methods for assessing cardiovascular status and compositions for use thereof
 JOURNAL Patent: US 6197505-A 128 06-MAR-2001;
 FEATURES Location/Qualifiers
 source 1..1856
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 41.0%; Score 824.6; DB 6; Length 1856;
 Best Local Similarity 98.3%; Pred. No. 4.4e-223;
 Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
 Qy 1 CCGGCCCGTTTGAAGAGCTACCCCGGACCAAGACCCGACACAGATACCTCCAGC 60
 Db 589 CCAGCCCGTTTGAAGAGCTACCCCGGACCAAGACCCGACACAGATACCTCCAGC 648
 Qy 61 TCCTCTCAACCCACCTTCCAGGTTGGAGACTTGGAGCTAAGCTTGTCTCCATCA 120
 Db 649 TCCTCTCAACCCACCTTCCAGGTTGGAGACTTGGAGCTAAGCTTGTCTCCATCA 708
 Qy 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGGCCCCCAGCCAGCTCCACATTAGAACAA 180
 Db 709 GGAATCTCCACCCAGAAATGGGTCTTTCTGGGCCCCCAGCCAGCTCCACATTAGAACAA 768

Qy 181 TGACAAATAGAGGGGAAATGGAATAAACAAGAGAAACGGTTTTCCAGACAGGGTT 240
 Db 769 TGACAAATAGAGGGGAAATGGAATAAACAAGAGAAACGGTTTTCCAGACAGGGTT 828
 Qy 241 TGGCCTACAGTCTGGAATGTGGGTAGCCATGCGCAAGTGTGAGGAGGCTGGCCGGGTG 300
 Db 829 TGGCCTACAGTCTGGAATGTGGGTAGCCATGCGCAAGTGTGAGGAGGCTGGCCGGGTG 888
 Qy 301 TGGTGGCTCATGCTCTAAATCCAGCACTTTGGGAGGCAAGGTGAGTACATCACTT 360
 Db 889 TGGTGGCTCATG-CTCTAAATCCAGCACTTTGGGAGGCAAGGTGAGTACATCACTT 947
 Qy 361 CCGGGAGTTTGGAGACAGGCTGGCCCAATCATGTTGAAACCCCATCTGTACTAAAAATACA 420
 Db 948 CCGGGAGTTTGGAGACAGGCTGGCCCAATCATGTTGAAACCCCATCTGTACTAAAAATACA 1007
 Qy 421 AAAGTTAGCTGGGCGTGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGAGCTGAGGCA 480
 Db 1008 AAAGTTAGCTGGGCGTGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGAGCTGAGGCA 1067
 Qy 481 TGAGAAATCGCTTGGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT 540
 Db 1068 TGAGAAATCGCTTGGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT 1127
 Qy 541 ACAAAAATTTAGTTGGATGTGGTGTGATGCTGTAGTCCAGCTACTTGGAGAGCTGGA 600
 Db 1128 ACAAAAATTTAGTTGGATGTGGTGTGATGCTGTAGTCCAGCTACTTGGAGAGCTGGA 1187
 Qy 601 GATGAAGGATTTGCTTGGAGCTGGGAGTCAAGGCTGAGGCTGAGGAGTGGCGCCACT 660
 Db 1188 GATGAAGGATTTGCTTGGAGCTGGGAGTCAAGGCTGAGGCTGAGGAGTGGCGCCACT 1247
 Qy 661 GCATCCAGCTTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAATAAAAAA 720
 Db 1248 GCATCCAGCTTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAATAAAAAA 1305
 Qy 721 GGAGGAGAGAGACTCAAGCAAGCCCTCACAGGACTGCTGAGGAGTGGCTGCTG 780
 Db 1306 GGAGGAGAGAGACTCAAGCAAGCCCTCACAGGACTGCTGAGGAGTGGCTGCTG 1365
 Qy 781 CAGCATGTGGCCCGCCAGGCGGGGACTCTGTAGCCACTGCTGGAGAGCACTCCCATCT 840
 Db 1366 CAGCATGTG--CCAGGCGGGGACTCTGTAGCCACTGCTGGAGAGCACTCCCATCT 1422
 Qy 841 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 891
 Db 1423 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 1473

RESULT 8
 BD231364
 LOCUS BD231364 1856 bp DNA linear PAT 17-JUL-2003
 DEFINITION Genes for assessing cardiovascular status and compositions for use thereof.
 ACCESSION BD231364
 VERSION BD231364.1 GI:33041134
 KEYWORDS JP 2002527079-A/128.
 SOURCE Homo sapiens (human).
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
 1 (bases 1 to 1856)
 Norberg,L.T., Andersson,M.K., Lindstrom,P.H.R. and Jonsson,L.
 Genes for assessing cardiovascular status and compositions for use thereof
 Patent: JP 2002527079-A 128 27-AUG-2002;
 PAIROSEAKENSINGU AB
 OS Homo sapiens (human)
 PN JP 2002527079-A/128
 PD 27-AUG-2002
 PF 13-OCT-1999 JP 2000576056
 PR 14-OCT-1998 US 60/104286,14-OCT-1998 US 60/104302 PI
 LEIF TORBUORN NORBERG, MARIA KRISTINA ANDERSSON, PER HARRY PI

RUTGER LINDSTROM,
PI LENA JONSSON
PC C1201/68.C12N15/09//G01N33/53.G01N33/566.C12N15/00 CC Genes
for assessing cardiovascular status
and compositions for
CC use thereof
FH Key Location/Qualifiers
FT source 1..1856
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1856
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 41.0%; Score 824.6; DB 6; Length 1856;
Best Local Similarity 98.3%; Pred. No. 4.4e-223;
Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
QY 1 CCGGCGGTTTGAAGAGCTCACCCCGACACAGGACCGGCACAGATACCTCCAGC 60
Db 589 CAGCCCGTTTGAAGAGCTCACCCCGACACAGGACCGGCACAGATACCTCCAGC 648
QY 61 TCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTCTCCATGA 120
Db 649 TCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTCTCCATGA 708
QY 121 GGAATCTCACCAGAAATGGGTCTTTCTGGCCCCCAGGCCAGCTCCCATTTAGAACAA 180
Db 709 GGAATCTCACCAGAAATGGGTCTTTCTGGCCCCCAGGCCAGCTCCCATTTAGAACAA 768
QY 181 TGACAAATAGAGGGGAATGGAAATTAACAGGAGAAACGGTTTCCAGGACAGGGTT 240
Db 769 TGACAAATAGAGGGGAATGGAAATTAACAGGAGAAACGGTTTCCAGGACAGGGTT 828
QY 241 TGSCCTTACAGTTGTGGATGTGGGTATGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 480
Db 829 TGSCCTTACAGTTGTGGATGTGGGTATGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 1067
QY 301 TGGTGGCTCATGCTCTAATCCAGCACTTTGGGAGGCCAAGGTGATGATCACTTTGAG 360
Db 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGGAGGCCAAGGTGATGATCACTTTGAG 947
QY 361 GCCGGAGTTTGAGACAGCTGGCCGACATGATGGTGAACCCCATCTGTACTAAAAATACA 420
Db 948 GCCGGAGTTTGAGACAGCTGGCCGACATGATGGTGAACCCCATCTGTACTAAAAATACA 1007
QY 421 AAAGTTAGCTGGGCGTGGTGTAGATGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 480
Db 1008 AAAGTTAGCTGGGCGTGGTGTAGATGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 1067
QY 481 TGAGATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTCAAAATAAAT 540
Db 1068 TGAGATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTCAAAATAAAT 1127
QY 541 ACAAAAAATTTAGTTGATGTGGTGTGATGCTGTAGTCTAGTCTAGTCTAGGAGCTGA 600
Db 1128 ACAAAAAATTTAGTTGATGTGGTGTGATGCTGTAGTCTAGTCTAGTCTAGGAGCTGA 1187
QY 601 GATGAAGGATTTGCTTGAAGCTGGGAGGTCAAGGCTGCAGTGCAGCGAGATGGCGCACT 660
Db 1188 GATGAAGGATTTGCTTGAAGCTGGGAGGTCAAGGCTGCAGTGCAGCGAGATGGCGCACT 1247
QY 661 GCATCTCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAATAAAAAAATAAAAA 720
Db 1248 GCATCTCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAATAAAAAAATAAAAA 1305
QY 721 GGAGAGGAGAGACTCAAGACAGCCCTTCAAGAGTCTGTGAGGCGCTTGCAGGTGTCTG 780
Db 1306 GGAGAGGAGAGACTCAAGACAGCCCTTCAAGAGTCTGTGAGGCGCTTGCAGGTGTCTG 1365
QY 781 CAGCATGTGGCCCCAGGCGGGGACTCTGTGAAGCCACTTGTGGAGAGCCACTCCCATCT 840

Db 1366 CAGCATGTG-CCAGGCGGGGACTCTGTAGCCACTGCTGGAGA-CCACTCCATCT 1422
QY 841 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGGTTT 891
Db 1423 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGGTTT 1473
RESULT 9
BD075256 1856 bp DNA linear PAT 27-AUG-2002
LOCUS Methods for assessing cardiovascular status and compositions for
DEFINITION use thereof.
ACCESSION BD075256
VERSION BD075256.1 GI:22620859
KEYWORDS JP 2001519660-A/129.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1856)
AUTHORS Norberg,L.T., Andersson,M.K. and Lindstrom,P.H.R.
TITLE Methods for assessing cardiovascular status and compositions for
use thereof
JOURNAL Patent: JP 2001519660-A 129 23-OCT-2001;
COMMENT EURONA MEDICAL AB
OS Homo sapiens (human)
PN JP 2001519660-A/129
PD 23-OCT-2001
PF 01-APR-1998 JP 1998542530
PI 04-APR-1997 US 60/042930
PI LEIF TORBJORN NORBERG,MARIA KRISTINA ANDERSSON,PER HARRY PI
RUTGER LINDSTROM
PC C1201/68.C07K14/72.C07K14/575.C12N9/48
CC ACE intron 16
FH Key Location/Qualifiers
FT source 1..1856
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1856
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 41.0%; Score 824.6; DB 6; Length 1856;
Best Local Similarity 98.3%; Pred. No. 4.4e-223;
Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
QY 1 CCGGCGGTTTGAAGAGCTCACCCCGACACAGGACCGGCACAGATACCTCCAGC 60
Db 589 CAGCCCGTTTGAAGAGCTCACCCCGACACAGGACCGGCACAGATACCTCCAGC 648
QY 61 TCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTCTCCATGA 120
Db 649 TCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTCTCCATGA 708
QY 121 GGAATCTCACCAGAAATGGGTCTTTCTGGCCCCCAGGCCAGCTCCCATTTAGAACAA 180
Db 709 GGAATCTCACCAGAAATGGGTCTTTCTGGCCCCCAGGCCAGCTCCCATTTAGAACAA 768
QY 181 TGACAAATAGAGGGGAATGGAAATTAACAGGAGAAACGGTTTCCAGGACAGGGTT 240
Db 769 TGACAAATAGAGGGGAATGGAAATTAACAGGAGAAACGGTTTCCAGGACAGGGTT 828
QY 241 TGSCCTTACAGTTGTGGATGTGGGTATGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 480
Db 829 TGSCCTTACAGTTGTGGATGTGGGTATGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 1067
QY 301 TGGTGGCTCATGCTCTAATCCAGCACTTTGGGAGGCCAAGGTGATGATCACTTTGAG 360
Db 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGGAGGCCAAGGTGATGATCACTTTGAG 947

```
QY 361 GCCGGAGTTTGGACCCAGCTGGCCACATGGTGAACCCCATCTGTACTATAAATACA 420
Db 948 GCCGGAGTTTGGACCCAGCTGGCCACATGGTGAACCCCATCTGTACTATAAATACA 1007
QY 421 AAAGTTAGTGGCGCTGGTGTAGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGCA 480
Db 1008 AAAGTTAGTGGCGCTGGTGTAGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGCA 1067
QY 481 TGAGAAATCGTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 540
Db 1068 TGAGAAATCGTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 1127
QY 541 ACAAAAATTTAGTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 600
Db 1128 ACAAAAATTTAGTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 1187
QY 601 GATGAAGATTTGCTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 660
Db 1188 GATGAAGATTTGCTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 1247
QY 661 GATGAAGATTTGCTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 720
Db 1248 GATGAAGATTTGCTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 1305
QY 721 GGAGAGGAGAGAGATCTAAGACAGCCCTCTACAGAGCTGCTGAGGCTGTAGGAGGCTGA 780
Db 1306 GGAGAGGAGAGAGATCTAAGACAGCCCTCTACAGAGCTGCTGAGGCTGTAGGAGGCTGA 1365
QY 781 GAGCATGTGGCCAGGCGGGGACTCTGTAGCCACTGCTGGAGGCACTCCCATCT 840
Db 1366 GAGCATGTGGCCAGGCGGGGACTCTGTAGCCACTGCTGGAGGCACTCCCATCT 1422
QY 841 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGTTT 891
Db 1423 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGTTT 1473
```

```
RESULT 10
BD075259 1856 bp DNA linear PAT 27-AUG-2002
LOCUS Methods for assessing cardiovascular status and compositions for
DEFINITION use thereof.
ACCESSION BD075259
VERSION BD075259.1 GI:22620862
KEYWORDS JP 2001519660-A/132.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Norberg, L.T., Andersson, M.K. and Lindstrom, P.H.R.
1 (bases 1 to 1856)
METHODS for assessing cardiovascular status and compositions for
use thereof
Patent: JP 2001519660-A 132 23-OCT-2001;
EUROA MEDICAL AB
OS Homo sapiens (human)
PN JP 2001519660-A/132
PD 23-OCT-2001
PF 01-APR-1998 JP 1998542530
PR 04-APR-1997 US 60/042930
PI LEIF TOREJORN NORBERG, MARIA KRISTINA ANDERSSON, PER HARRY PI
RUTGER LINDSTROM
PC C1201/68, C07K14/72, C07K14/575, C12N9/48
CC Deletion
FH Key Location/Qualifiers
FT source 1. .1856
FT Location/Qualifiers
1. .1856 /organism="Homo sapiens"
/ncbi_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
ORIGIN
```

```
Query Match 41.0%; Score 824.6; DB 6; Length 1856;
Best Local Similarity 98.3%; Pred. No. 4.e-223;
Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
QY 1 CGGCGCGCTTTGAAAGAGCTCACCCCGCACACAGGACCCCGCACACATACCTCCAGC 60
Db 589 CGAGCGCGCTTTGAAAGAGCTCACCCCGCACACAGGACCCCGCACACATACCTCCAGC 648
QY 61 TCCCTCTCAACCCACCCCTTTCCAGGTTGGAGAACTTGAGGCATAAAATTGCTTCCATGA 120
Db 649 TCCCTCTCAACCCACCCCTTTCCAGGTTGGAGAACTTGAGGCATAAAATTGCTTCCATGA 708
QY 121 GGAATCTCCACCCAGAAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCAATTAGAACAA 180
Db 709 GGAATCTCCACCCAGAAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCAATTAGAACAA 768
QY 181 TGACAAATAGAGGGGAAATGGAATAAAGAGGAAACGGTTTCCAGGACAGAGGTT 240
Db 769 TGACAAATAGAGGGGAAATGGAATAAAGAGGAAACGGTTTCCAGGACAGAGGTT 828
QY 241 TGGCCCTCAAGTTGGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCCGGTG 300
Db 829 TGGCCCTCAAGTTGGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCCGGTG 888
QY 301 TGGTGGCTCATGCTCTTAATCCAGCACTTTGGGGAGGCCAAGGTGAGTAGATCACTTGAG 360
Db 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGGGAGGCCAAGGTGAGTAGATCACTTGAG 947
QY 361 GCCGGAGTTTGGAGACCCAGCTGGCCCAACATGGTGAACCCCATCTGTACTAAAAATACA 420
Db 948 GCCGGAGTTTGGAGACCCAGCTGGCCCAACATGGTGAACCCCATCTGTACTAAAAATACA 1007
QY 421 AAAGTTAGTGGCGCTGGTGTAGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCA 480
Db 1008 AAAGTTAGTGGCGCTGGTGTAGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCA 1067
QY 481 TGAGAAATCGTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 540
Db 1068 TGAGAAATCGTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 1127
QY 541 ACAAAAATTTAGTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 600
Db 1128 ACAAAAATTTAGTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 1187
QY 601 GATGAAGATTTGCTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 660
Db 1188 GATGAAGATTTGCTTGGATGTGTGTGATGCTGTAGTCTGTAGTCTGTAGGAGGCTGA 1247
QY 661 GCATCCAGCTGGGCAACAGAGTGAACCCCTGTCTCAGAAAAAATAAATAAATAA 720
Db 1248 GCATCCAGCTGGGCAACAGAGTGAACCCCTGTCTCAG--AAAGAAAAAATAAATAAATAA 1305
QY 721 GGAGAGGAGAGAGACTCAAGCAGCCCTCTCAGAGCACTGCTGAGGCGCTGCGAGTGTCTG 780
Db 1306 GGAGAGGAGAGAGACTCAAGCAGCCCTCTCAGAGCACTGCTGAGGCGCTGCGAGTGTCTG 1365
QY 781 CAGCATGTGGCCAGGCGGGGACTCTGTAGCCACTGCTGGAGGCACTCCCATCT 840
Db 1423 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGTTT 1473
RESULT 11
HSATICE
LOCUS H.sapiens angiotensin I-converting enzyme gene (intron 16).
DEFINITION H.sapiens angiotensin I-converting enzyme gene (intron 16).
ACCESSION X62855
VERSION X62855.1 GI:28921
KEYWORDS Alu-like repetitive sequence; angiotensin I-converting enzyme;
```

polymorphic variation.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1856)
 AUTHORS Rigat,B., Hubert,C., Corvol,P. and Soubrier,F.
 TITLE PCR detection of the insertion/deletion polymorphism of the human
 angiotensin converting enzyme gene (ACE1) (dipeptidyl
 carboxypeptidase 1)
 JOURNAL Nucleic Acids Res. 20 (6), 1433 (1992)
 MEDLINE 92220641
 PUBMED 1313972
 REFERENCE 2 (bases 1 to 1856)
 AUTHORS Soubrier,F.
 TITLE Direct Submission
 JOURNAL Submitted (25-OCT-1991) F. Soubrier, Inserm Unit 36 College de
 France, 3 Rue d'Ulm, 75005 Paris, FRANCE
 COMMENT Related sequences: Hubert,C. et al, J.Biol.Chem. 266:15377-15383
 (1991).
 FEATURES
 Location/Qualifiers
 1..1856
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="q23"
 /clone="Lambda 311"
 /tissue_type="placenta"
 /clone_lib="genomic"
 880..1105
 /note="incomplete Alu type sequence"
 repeat_region 1106..1301
 /note="incomplete Alu type sequence"
 repeat_region 1451..1738
 /note="insertion/deletion polymorphism due to
 presence/absence of sequence"
 complement(1451..1738)
 /note="Alu type sequence in antisense orientation"
 ORIGIN
 Query Match 41.0%; Score 824.6; DB 9; Length 1856;
 Best Local Similarity 98.3%; Pred. No. 4.4e-223;
 Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
 QY 1 CGGGCCCGTTGAAGAGCTCACCCCGACACAGACCGCACAGATACCTCCAC 60
 Db 589 CCAGCCCGTTGAAGAGCTCACCCCGACACAGACCGCACAGATACCTCCAC 648
 QY 61 TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAATCTTCCATCA 120
 Db 649 TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAATCTTCCATCA 708
 QY 121 GGAATCTCAACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCACTAGAACAA 180
 Db 709 GGAATCTCAACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCACTAGAACAA 768
 QY 181 TCACAAATAGAGGGGAAATGGAATAACAGGAGAAACGGTTTCCAGGACAGGGTT 240
 Db 769 TCACAAATAGAGGGGAAATGGAATAACAGGAGAAACGGTTTCCAGGACAGGGTT 828
 QY 241 TGGCTTACAAAGTTGTGGATGTGGGTATCCCATGCCAAGTGTGAGGGAGGCTGGCCGGGTG 300
 Db 829 TGGCTTACAAAGTTGTGGATGTGGGTATCCCATGCCAAGTGTGAGGGAGGCTGGCCGGGTG 888
 QY 301 TGGTGGCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGAG 360
 Db 889 TGGTGGCTCATG-CTTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGAG 947
 QY 361 GCGGGAGTTTGAGACCGCTGGCCCAACATGGTGAACCCCATCTGTACTAAAAATACA 420
 Db 948 GCGGGAGTTTGAGACCGCTGGCCCAACATGGTGAACCCCATCTGTACTAAAAATACA 1007

QY 421 AAAGTTAGCTGGCGGTGGTGGTAGATGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 480
 Db 1008 AAAGTTAGCTGGCGGTGGTGGTAGATGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 1067
 QY 481 TGAGAATCGCTTTGAGCCCGAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT 540
 Db 1068 TGAGAATCGCTTTGAGCCCGAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT 1127
 QY 541 ACACAAAATTTAGTGGATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
 Db 1128 ACACAAAATTTAGTGGATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1187
 QY 601 GATGAAGGATTTGCTTGAGCTGGGAGGTCAAGGCTGAGTGGAGCCGAGATGGCGCACT 660
 Db 1188 GATGAAGGATTTGCTTGAGCTGGGAGGTCAAGGCTGAGTGGAGCCGAGATGGCGCACT 1247
 QY 661 GCATCCAGCTGGGCAACAGAGTGGAGCCCTGTCTCAGAAAAAATAAAAAA 720
 Db 1248 GCATCCAGCTGGGCAACAGAGTGGAGCCCTGTCTCAG--AAAGAAAAAATAAAAAA 1305
 QY 721 GGAGGAGGAGAGACTCAAGCAGCGCCCTCACAGGACTGCTGAGGCCCTGCGAGTGTCTG 780
 Db 1306 GGAGGAGGAGAGACTCAAGCAGCGCCCTCACAGGACTGCTGAGGCCCTGCGAGTGTCTG 1365
 QY 781 CAGCATGTGGCCCGGAGGCTCTCTAAGCCACTGCTGGAGGACCACTCCCATCT 840
 Db 1366 CAGCATGTG--CCAGGCGGAGGCTCTCTAAGCCACTGCTGGAGGACCACTCCCATCT 1422
 QY 841 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 891
 Db 1423 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 1473
 RESULT 12
 AC118572
 LOCUS Lemur catta clone LB2-253K19, WORKING DRAFT SEQUENCE, 12 ordered
 DEFINITION pieces.
 ACCESSION AC118572
 VERSION AC118572.2 GI:32526927
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
 SOURCE Lemur catta (ring-tailed lemur)
 ORGANISM Lemur catta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Strepsirhini; Lemnidae; Lemur.
 REFERENCE 1 (bases 1 to 195863)
 AUTHORS Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
 Peng,Z., Malinov,I. and Rubin,E.M.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 195863)
 AUTHORS Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and
 Cheng,J.-F.
 TITLE Direct Submission
 JOURNAL Submitted (19-APR-2002) Genome Sciences, Lawrence Berkeley National
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
 REFERENCE 3 (bases 1 to 195863)
 AUTHORS Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
 Peng,Z., Malinov,I. and Rubin,E.M.
 TITLE Direct Submission
 JOURNAL Submitted (11-JUL-2003) Genome Sciences, Lawrence Berkeley National
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
 COMMENT On Jul 11, 2003 this sequence version replaced gi:20198536.

Sequence Produced by Berkeley PGA

Web site: <http://pga.lbl.gov>

Center Code: PGABERK

Center Project Name: L027

Bac Clone Name: LB2-253K19

This sequence has been compared to sequences of other species
 using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be
 viewed at:

http://pga.lbl.gov/cgi-bin/search_cvcgd?type=nkvalue=ACE

The order-orientation of the draft sequence was accomplished by using:

Avid (<http://baboon.math.berkeley.edu/mavid/>),
Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

If the Bac Library Name is LB1 to LB4, please see website for the description: <http://www-gsd.lbl.gov/cheng/BAC.html>
These libraries are available through the BACPAC Resources Center: <http://www.chori.org/bacpac/libraryres.htm> as LBNL-1 to LBNL-4.

Summary Statistics:

Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 12 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 10817: contig of 10817 bp in length

* 10818 10917: gap of unknown length

* 10918 19222: contig of 8305 bp in length

* 19223 19322: gap of unknown length

* 19323 52285: contig of 32963 bp in length

* 52286 52385: gap of unknown length

* 52386 75793: contig of 23408 bp in length

* 75794 75893: gap of unknown length

* 75894 99277: contig of 23384 bp in length

* 99278 99377: gap of unknown length

* 99378 107543: contig of 8166 bp in length

* 107544 107643: gap of unknown length

* 107644 118868: contig of 11225 bp in length

* 118869 121200: gap of unknown length

* 121201 121300: gap of unknown length

* 121301 139834: contig of 18534 bp in length

* 139835 139934: gap of unknown length

* 139935 156316: contig of 16382 bp in length

* 156317 156416: gap of unknown length

* 156417 167012: contig of 10596 bp in length

* 167013 167112: gap of unknown length

* 167113 195863: contig of 28751 bp in length.

FEATURES

source

```
1..195863
/organism="Lemur catta"
/mol_type="genomic DNA"
/db_xref="taxon:9447"
/clone="LB2-253K19"
```

ORIGIN

```
Query Match      24.4%; Score 490.6; DB 2; Length 195863;
Best Local Similarity 72.0%; Pred. No. 9.3e-128;
Matches 860; Conservative 0; Mismatches 264; Indels 71; Gaps 14;

QY 843 CTCCTCATTTCTAGACCTGCTGCTATACAGTCACCTTTATGTGGTTTCGCCAATTITA 902
DB 128866 CCCCAGCTGCTGGCGCTTCTCCAGGGGCGAGAACTTCCAGTGGTCTCTGC-ATTGA 128924
QY 903 TTCCAGCTCTGAATAATCTCTAGCTCCCTTTACAGCGAG-----AGGTGAGCTAAGGGCT 957
DB 128925 CTCCTGCTCTGAGAGCTGTGAGTCCCTGCAAGAGCGAGGTGGTGGCTGAGCCAAAGGGCT 128984
QY 958 GGAGCTCAAGGATTCMAACCC-----CTACAGATCTCAGAAATGTGATGCCACG 1009
DB 128985 GGGGCTCAGCTGTTCGACCCCAACCCCAACCCCAACCTCCAGATCTGACAGCGTGTGCGCACA 129044
```

```
QY 1010 TCCCGAAATATGAAGACCTGTATTGGGCATGGGAGGCTGGCGAGACAAGCGGGGAGA 1069
DB 129045 TCCCGAAGTACGAAAGAACTGTATTGGGCTTGAAGGGCTGGCGAGACCAAGTGGGAAGA 129104
QY 1070 GCCATCTCTCCAGTTTATCCCGAAATACGTGAAACTCATCAACCAAGGCTGCCCGGCTCAAT 1129
DB 129105 GCCATCTCTCTCTTTTCCCAAAATACGTGAAACTCATCAACCAAGGCTGCCCTGCTCAAC 129164
QY 1130 GGTGAGTCTCTGCTGCCCAACATCACTGCACTTGGGTCCCTTCATTTTCTCAAGAGGT 1189
DB 129165 GGTGAGTCTCCAGTAGCC-CTGGCTCTCTGGTCCCTTGGCGCCCTCGGGTTCTCTCAAGAGGT 129223
QY 1190 GCTGTGAACCCCAAGCCCTAGGAAAGGTAGATCCCTGGAGGAGCGAGTAA----- 1241
DB 129224 CCTCTGAGACCGTAAGTCTGGGGAAGG-AGGGCCCTGGAGCGGCGAGGAGGTGAGGG 129282
QY 1242 ----TGTTGGTGTGGGAGAGCTGGCTGTGTCCCTCTGTAG-GTATGTAGATCAGAGG 1296
DB 129283 ATGGTCAGGTGTTCGAAGAGACTGGCTGTGTCTCCCTTCGAGAGCTACGACAGCGAGGG 129342
QY 1297 GACTCGTGGAGTCTATGTACGAGACACCATCCCTGGAGCAAGACCTCGAGCGGCTCTTC 1356
DB 129343 GATTCTGGAGGTCTATGTACGAGACCATCCCTGGAGCAAGACCTCGAGCGAGCTCTAC 129402
QY 1357 CAGGAGCTGCAGCCACTCTACCTCAACCTGCATGCTACGTGCGCGCGGCGCTGCACCGT 1416
DB 129403 CAGGAGCTGCAGCGCTTTTACCTGAACCTGCACGCTACGTGCGCGCGGCGCTGCACCG 129462
QY 1417 CACTACGCGGCGCCAGCACATCAACTGAGGGGCGCCATCTCTGTCTACCTGCTGGGTAG 1476
DB 129463 CACTACGCGGCGCCAGCACATCAACTGAGGGGCGCCATCTCTGTCTCATCTGCTGGGTAG 129522
QY 1477 GGCACATGTCGGGCTTTCAGGAGGTGAAGACGACACAGTGTGAGTGAGGTTGGGAC 1536
DB 129523 GGCACAGCTGGGCTTTCAGGAGGTGAAGACG-CCAAGTAGCGGTGAGGCGCGGAC 129581
QY 1537 AGGCTGACTAGAGGTAGGAGCAGGCTGGGGACTGAGAGACTCCAGCCCTGTGGGGA 1596
DB 129582 AGGCTGSCCAGAGGGTGCAGAGCAGGCGCGTGTGGGAAGACTCGAGCCCTGGGGGA 129641
QY 1597 TGGTTCGCCAGCTGGAGGGGGTGGGCTGTGGAGTGGGGAGCCCCCAGCTTGATCTG 1656
DB 129642 AGGTGGCCAAAGCTGCAGAGGGTGGGCTGGGGGGG-----CTGCTTGCAATTG 129694
QY 1657 GTGCCACATTCACATGACATCTATGTCGGGCAAGTCACATGGA-----TG 1702
DB 129695 ATGCCACATTTGCTACATCCACATGTGCAAGTCACATGGAAGATGCCCCCATGGTG 129754
QY 1703 GGGGAAGAAAGTTAATAATCTTTGTCAGGAGACCCAGGCACCCACATCAACCAATTTGTGAT 1762
DB 129755 GGGCGAGAGGTTTCAATAATCTTGTGAGGAGAGACCACTACCTGTCTATACTTGAAGTG 129814
QY 1763 CTTAGAGGGCAGAGAGAGGCTGTGAGTGGAGCTGGGAGGCTTTGCCAAGAGTGGCC 1822
DB 129815 CTTTGGTCCAGGAGAGCGG-----GGGAGGCGCTCCCAAGGGGCGGCG 129858
QY 1823 TGTGAGCAGGCGCTCGGAAGATGACAGGCTTTGACAGATGGGAAGTGGGGATGAGAGA 1882
DB 129859 TGTGAG--TGCCCTTGAAGGGCGGAGGTTTGTTCAGAGGGGAATGAGGAGGAGAGA 129916
QY 1883 CAGACGAGTGTTCAGGCGCAAGGGAACTGGAAACAAAGAAGAACCT-GAGAATGTAAATCT 1941
DB 129917 CAGGCACAGCATTCAGGCGGAGGAACTGGGGCAAGAAGAATCTCGAGAAATGTAGATCC 129976
QY 1942 ACTTCAAC-CCTTGAACCTCTTTGCCAAGGGCTCAATCTCAGATGCCCTGAAT 1995
DB 129977 ACCTCAACTTGGGAGCCCTCAGTCCAGGAGTTGCCAACTCCAATGCCCGAAT 130031
```

RESULT 13

AL596246

LOCUS

144136 bp DNA linear

ROD 21-DEC-2002

FEATURES	a paired query at least 30.				
source	Location/Qualifiers				
	1..144136				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
	/chromosomes="11"				
	/clone="RP23-186E14"				
	/clone_lib="RPC1-23"				
ORIGIN					
Query Match	14.9%; Score 299.8; DB 10;	Length 144136;			
Best Local Similarity	65.4%; Pred. No. 1.8e-73;				
Matches 576;	Conservative 0;	Mismatches 272;	Indels 33;	Gaps 8;	
Qy	774 GTGCTGGCAGCATGTGGCCCGGCCAGCGGCAGCTCTGTAGCACACTGCTGGAGAGCACTC	833			
Dd	139203 GTGTGTGCAGCATGTACACCCCAG-TGGCACCCCTAGGAGTCATTCGACCGCTGGCTT	139261			
Qy	834 CCATCTTTTCTCCATTCTCTAGACCTGTCGCTAT-ACAGTCACTTTTATGTGGTTTC	892			
Dd	139262 CCCCATTCGCCCTACCCCACTAGACTTCTGCCGCTTTTGTCATTTCCAAGTCTCTT	139321			
Qy	893 GCCAAATTTATTCAGCTCTGAATTTCTCTGAGCTCCC-CTTAGCAGCAGAGGTGAGCTA	951			

Penestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grallan, J., Grand-Pierre, N., Grant, G., Hagos, B., Hearford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Landers, T., Largocque, K., Lehec, J., Levine, R.,
Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
McGurk, A., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Olivari, T.M.,
Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C.,
Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A.,
Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zimmer, A. and
Zody, W.

Direct Submission
Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 217088)

Barren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Archchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
FitzGerald, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Roh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasany, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (11-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 11, 2003 this sequence version replaced gi:11225431.
All repeats were identified using RepeatMasker:
[http://ftp.genome.washington.edu/RM/RepeatMasker.html]

----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L3199
Center Clone name: 321 P 3

```

----- Summary Statistics -----
Center: Clone name: 2411-3
Sequencing vector: M13; 7M7815; 34% of reads
Sequencing vector: Plasmid; n/a; 66% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 215379 bases at least Q40
Consensus quality: 215982 bases at least Q50
Consensus quality: 216295 bases at least Q30
Insert size: 213000; agarose-fp
Insert size: 216588; sum-of-contigs
Quality coverage: 13.3 in Q20 bases; agarose-fp
Quality coverage: 13.1 in Q20.

```

* quality coverage: 13.1 in Q20.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
 * be preserved.

```

1 68732: contig of 68732 bp in length
* 68733 68832: gap of 100 bp
* 68833 74687: contig of 5855 bp in length
* 74688 74787: gap of 100 bp
* 74788 83104: contig of 8317 bp in length
* 83105 83204: gap of 100 bp
* 83205 91107: contig of 7903 bp in length
* 91108 91208: gap of 100 bp
* 91209 129624: contig of 38417 bp in length
* 129625 129724: gap of 100 bp
* 129725 217088: contig of 87364 bp in length.
FEATURES
      source
      1..217088
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      /chromosome="11"
      /map="11"
      /clone="RP23-321E3"
      /clone_lib="RPC1-23 Female Mouse BAC"
misc_feature
      1..68732
      /note="assembly_fragment"
      clone_end:SP6
      vector_side:left
misc_feature
      68833..74687
      /note="assembly_fragment"
      74788..83104
      /note="assembly_fragment"
      83205..91107
      /note="assembly_fragment"
      91208..129624
      /note="assembly_fragment"
      129725..217088
      /note="assembly_fragment"
      clone_end:T7
      vector_side:right
ORIGIN
Query Match          14.9%; Score 299.8; DB 2; Length 217088;
Best Local Similarity 65.4%; Pred. No. 1.9e-73;
Matches 576; Conservative 0; Mismatches 272; Indels 33; Gaps 8;

QY      774  GTGCTGCGAGCATGTGGCCCGGAGCTCTGTAAGCCACTGCTGGAGAGCCACTC  833
Db      774  |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      55062 GTGTGTGCGAGCATGTACCCGAGG-TGGGCACCTTAGGAGCTCATTCGACCGCTGCTC  55004
Db      55062 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      834  CCATCCTTTCTCCCATTTCTCTAGACCTGTGCTGCTAT-ACAGTCACATTTTATGTGTTTC  892
Db      834  |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      55003 CCCCATTCCCTCTACCCGACCTAGACTTCTTGCCTGCTCTTGTCTATTCTCAAGTGTCTTT  54944
Db      55003 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      893  GCCAATTTTATTCAGCTCTCAAAATCTCTGAGTCTCC-CTTACAAGCAGAGGTGAGCTA  951
Db      893  |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      54943 ACCAGTTTATTCCACTCTGAACTCAATGACCAAGTTCCTCAATAGCAGATGAGAG  54884
Db      54943 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      952  AGGGCTGGAGCTCAAGGCATTCAAAACCCCTA---CCAGATCTGACGAATGTGATGGCCAC  1008
Db      952  |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      54883 TGAACCAAGGCTCAAGGCATCTGACCCCTTCATTCCAGATCTAAACAAACATGATGGCCAC  54824
Db      54883 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      1009 GTCCCGGAAATATGAGACCTGTTATGGCGATGGAGGGCTGGCGAGACAGAGCGGGAG  1068
Db      1009 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      54823 ATCCCGGAAATATGAAGAATTGCTATGGCGATGGAAGAGCTGGAGAGACAGAGTGGGGAG  54764
Db      54823 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      1069 AGCCATCTCTCCAGTTTATACCGGAAATACGTGGAACTCATCAACACAGGCTCCCGGCTCAA  1128
Db      1069 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      54763 AGCCATCTCTCTTTTTCCAAAGTATGTGGAGTTCTTCCAAACAAAGATTGCCAAGCTCAA  54704
Db      54763 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      1129 TGGTAGTCCCTGTGCGCAAATCACTGGACATGGGTCCTCC---TTCAATTTCTCTCAA  1184
Db      1129 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      54703 TGGTAGTACCAATGTGTGGTAGCTCTGCTGGCTGGCTGCTCCCACTCACTCAACTCTCTCA  54644
Db      54703 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||
QY      1185 GAGGTGCTGTGAAACCCCAAGCCTAGGAAAAAGGTAGATCCCTCGAGAGAGGCAGGTA----  1240
Db      1185 |||||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||G|||

```


b	54643	
---	-------	--

Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,V., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwuonu,G., Olarpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Ffankoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,B., Song,X.-Z., Sorelle,R., Taylor,C., Teimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 213169)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 213169)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GFQE
Center clone name: CH230-44G5
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 195455 bases at least Q40
Consensus quality: 199513 bases at least Q30
Consensus quality: 201879 bases at least Q20
Estimated insert size: 207657; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

* 1 209047: contig of 209047 bp in length
* 209048 209147: gap of unknown length
* 209148 210339: contig of 1492 bp in length
* 210640 210739: gap of unknown length
* 210740 213169: contig of 2430 bp in length.

FEATURES
source

```

source
1. .231369
/organism="Rattus norvegicus"
/mol_type="genomic"
/db_xref="taxon:10116"
/clone="CH230-4465"
1. .1473
/misc_feature
1. .8737
/misc_feature
1. .8737
/misc_feature
1. .145720
/misc_feature
1. .209047
/misc_feature
1. .209047

```

ORIGIN

Query Match	13.8%	Score 277.2	DB 2	Length 213169
Best Local Similarity	67.3%	Pred. No. 5e-67		
Matches 479	Conservative 0	Mismatches 203	Indels 30	Gaps 5
QY	940	AGAGGTGAGCTAAGGGCTGGAGCTCAAGGCATTCAAAACC---CTACCAGATCTGACGAA	996	
Db	53506	AGCAGTGAAGCAAGAACTGAGGCTCAAACTTCGACCCCTTCTTCACGATCTGACAAA	53447	
QY	997	TGTGATGGCCACGTCCTCCGGAATATGAGACCTGTTATGGGCATGGGAGGGCTGGCGAGA	1056	
Db	53446	TATTAATGGCCACGTCCTCCGGAATACGAAGAAATGCTTTTGGGTGTGGAAGAGCTGGCGAGA	53387	
QY	1057	CAAGCGGGGAGAGCCATCTCCAGTTTTACCGAAATAGCTGGAACTCATCAACACGAC	1116	
Db	53386	CAAGTGGGAGAGCCATCTCTCCCTTTTCCCAAAGTACGTGGACTTCTTCCAAACAGAT	53327	
QY	1117	TGCCCGCTCAATGGTGAAGTCCCTGCTGCCAACATCATCTGGGCACCTTGGGTCC---CTTC	1172	
Db	53326	CGCCAAAGCTCAACGGTGAAGTTACCACTGCTGGTACCTCTGGCTGGCTGCTCCACGACCC	53267	
QY	1173	ATTTTCTCAAGAGGTGCTGTGAACCCCAAGCCTAGGAAAGGTAGATCCCTGGAGGA	1232	
Db	53266	TCAAACTCTCAAGATCTCTGGGATCTCAGGCTAGGGCAAGTAGGGACCTAGAGGT	53207	
QY	1233	GGCAGGTA-----ATGTGGTGTGGAGAGCCCTGGCTGTGTCCCTCTGTGA	1278	
Db	53206	GTCAAGGATATATGGGTGAGTGAACGGTTGTGGAGAGACTTGGCTGTCTCTGTGCA	53147	
QY	1279	GGCTATGTAGATGCAAGGACCTGCTGGAGGTCTATGTACGAGACACCATCCCTGGAGCAA	1338	
Db	53146	GGCTACTCTGATGCAAGGAGTTCCTGGAGATCCTCATATGAGTCCGACGCTTGGAGCAA	53087	
QY	1339	GACCTGGAGGGCTCTTCAGAGAGCTGCAGCCACTCTACCTCAACCTGCATGCTACGTG	1398	
Db	53086	GACCTGAAAACCTATACAGAGAGCTGCAGCGCTCTACCTCAACCTGCATGCTATGTG	53027	
QY	1399	CGCCGGGCTTGCACCGTCACTACGGGGCCAGACATCAACCTGGAGGGGCCATTTCCT	1458	
Db	53026	CGCCGCTCCCTGCACCGCCATTATGGGTCTGAGTACATCAACCTGGATGGTTCCTTCCT	52967	
QY	1459	GCTCACTCTGGGTAAAGGACATGTCGGCCCTTGAGAGGGTAAAGACGACCAAGT	1518	
Db	52966	GCCCACTCTAGTTAAGACATCCGTCCAGCTTCAAGAGGGTGGCAATGGA-----	52914	
QY	1519	GTGAGTGAAGGTGGGACAGGGCTGACTAGAGGGTAGGGAGC-AGGCTGGGGCTGAGAG	1577	
Db	52913	-TGATAAAGGAACAAGATAGTCTGGCCACAGGGCAGAGATCAAGCCAGAAATAGTGACAG	52855	
QY	1578	ACTCCAGCCCTGTGGGGAGTGTGCCAGGCTGGAGGGGGGTGGGGCTGG	1629	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:20 ; Search time 1233.75 Seconds
(without alignments)
11849.778 Million cell updates/sec

Title: US-10-017-724-1_COPY_3000_5784

Perfect score: 2785

Sequence: 1 gaggacttgacgtgtgaggg.....aattgtaaaaggtttctt 2785

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2785	100.0	5784	6	ABL62660 Colon ade
2	2785	100.0	5784	6	Abz35081 Human gen
3	2785	100.0	5784	6	Abk10277 Human cDN
4	2785	100.0	5784	6	Abn95662 Gene #216
5	2785	100.0	5784	6	Aad40611 Human thr
6	2785	100.0	5784	6	Abv94768 Human pan
7	2785	100.0	5784	8	Abq77406 Human THB
8	2785	100.0	5784	9	Aad55115 Human THB
9	2785	100.0	5784	10	Add31094 Human cDN
10	2785	100.0	5784	10	Ade85079 Farnesyl
11	2785	100.0	5784	12	Adq18777 Human sof
12	2745.4	98.6	5843	12	Adq23188 Human sof
13	2735.8	98.2	3787	3	Aac77795 Human can
14	2709.6	97.3	5696	9	Ach03798 Human cDN
15	2709.6	97.3	5696	10	Adj56219 Bovine cD
16	2456.2	88.2	3522	4	Aah34535 Human col
17	1694	60.8	2780	12	AcF57539 Human thr
18	809.4	29.1	3596	3	Aaa88669 Human thr
19	809.4	29.1	3596	10	Abx14061 cDNA enco
20	766.6	27.5	4170	6	Ab135019 Rat cDNA
21	753.6	27.1	4108	2	Aaz32016 Human MET

22	753.6	27.1	4108	5	AAC90073	Aac90073 L07803 cD
23	651.8	23.4	713	4	AAH34534	Aah34534 Human col
c 24	595.6	21.4	619	3	AAH78040	Aaz78040 cDNA enco
c 25	595.6	21.4	619	4	AAI28778	Aai28778 Colon tum
c 26	595.6	21.4	619	8	ABZ32964	Abz32964 Human col
c 27	594.8	21.4	848	8	AAZ17196	Aaz17196 Human gen
c 28	540.2	19.4	637	4	AA558772	Aas58772 cDNA #144
c 29	540	19.4	540	6	ABV98659	Abv98659 Human pan
c 30	516	18.5	516	6	ABK45682	Abk45682 cDNA enco
c 31	515	18.5	551	6	ABV98694	Abv98694 Human pan
c 32	497.4	17.9	570	6	ABV99051	Abv99051 Human pan
c 33	496.8	17.8	529	6	ABV97469	Abv97469 Human pan
c 34	495.4	17.8	551	6	ABV98796	Abv98796 Human pan
c 35	495.4	17.8	5289	2	AZ32052	Aaz32052 Human MET
c 36	495.4	17.8	5289	5	AAC90309	Aac90309 AB005287
c 37	491.8	17.7	7231	6	ABS62773	Abs62773 Prostate
c 38	491.8	17.7	7231	10	ADJ56211	Adj56211 Murine cD
c 39	491.8	17.7	7232	6	AAS94912	Aas94912 Human DNA
c 40	490.2	17.6	3967	11	ADN39634	Adn39634 Cancer/an
c 41	490.2	17.6	4433	2	AZ32014	Aaz32014 Human MET
c 42	490.2	17.6	4434	5	AAC90071	Aac90071 X04665 cD
c 43	490.2	17.6	4434	10	ADD18445	Add18445 Human pro
c 44	490.2	17.6	5721	4	AAH02923	Aah02923 Human she
c 45	490.2	17.6	5722	2	AZ32012	Aaz32012 Human MET

ALIGNMENTS

RESULT 1

ABL62660

ID ABL62660 standard; DNA; 5784 BP.

XX

AC ABL62660;

XX

DT 15-MAY-2002 (first entry)

XX

XX Colon adenocarcinoma related gene sequence SEQ ID NO:997.

DE

XX

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

XX gene; ds.

XX

OS Homo sapiens.

XX

PN WO200194629-A2.

XX

PD 13-DEC-2001.

XX

PF 30-MAY-2001; 2001WO-US010838.

XX

PR 05-JUN-2000; 2000US-0209473P.

PR

PR 05-JUN-2000; 2000US-0209531P.

PR

PR 18-SEP-2000; 2000US-0233133P.

PR

PR 18-SEP-2000; 2000US-0233617P.

PR

PR 20-SEP-2000; 2000US-0234009P.

PR

PR 20-SEP-2000; 2000US-0234034P.

PR

PR 20-SEP-2000; 2000US-0234052P.

PR

PR 22-SEP-2000; 2000US-0234509P.

PR

PR 22-SEP-2000; 2000US-0234567P.

PR

PR 25-SEP-2000; 2000US-0234923P.

PR

PR 25-SEP-2000; 2000US-0234924P.

PR

PR 25-SEP-2000; 2000US-0235077P.

PR

PR 25-SEP-2000; 2000US-0235082P.

PR

PR 25-SEP-2000; 2000US-0235134P.

PR

PR 25-SEP-2000; 2000US-0235280P.

PR

PR 26-SEP-2000; 2000US-0235637P.

PR

PR 26-SEP-2000; 2000US-0235638P.

PR

PR 27-SEP-2000; 2000US-0235711P.

PR

PR 27-SEP-2000; 2000US-0235720P.

PR

PR 27-SEP-2000; 2000US-0235840P.

PR

PR 27-SEP-2000; 2000US-0235863P.

PR

PR 28-SEP-2000; 2000US-023602BP.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Zi;
 PI
 DR WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 XX Claim 1; SEQ ID NO 997; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (II) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX
 SQ Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;
 Query Match 100.0%; Score 2785; DB 6; Length 5784;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGGACTTGGACGGTGATGGACGGGGTGATATTTGTAAGATGATTTTGACAATGACAAC 60
 DB 3000 GAGGACTTGGACGGTGATGGACGGGGTGATATTTGTAAGATGATTTTGACAATGACAAC 3059
 QY 61 ATCCAGAGATTTGATGATGTGTCTCTGAAACCAATGCGATCAGTGAGACAGACTTCAGG 120
 DB 3060 ATCCAGAGATTTGATGATGTGTCTCTGAAACCAATGCGATCAGTGAGACAGACTTCAGG 3119
 QY 121 AACTTCCAGATGGTCCCTTGGATCCCAAGGACCAACCAAAATGATCCCAACTGGGTC 180
 DB 3120 AACTTCCAGATGGTCCCTTGGATCCCAAGGACCAACCAAAATGATCCCAACTGGGTC 3179
 QY 181 ATTCCGCATCAGGACGAGTGTGTTGACAGGCAACTCGGACCCCGGATCCCTGTA 240

small airway epithelium, renal epithelium, umbilical artery smooth muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle, dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes, aortic smooth muscle, mesangial cells, coronary artery smooth muscle, bronchial smooth muscle, uterine smooth muscle, lung fibroblast, osteoblasts or prostate stromal cell. The gene expression profile is used for determining the level of RNA expression for a sample, determining the phenotype of a cell and distinguishing cell types. The gene or a protein expression profile is useful in identifying disease pathologies involving alterations of gene expression. The assessment of expression profiles may provide meaningful information with respect to tumour type and stage, treatment methods, and prognosis. The gene or protein expression profile may also be used for creating microarrays. The microarray is useful for genetic and physical mapping of genomes, DNA sequencing, genetic or medical diagnosis, genotyping of organisms, confirming cell or tissue identifications and in identifying promising antibiotics, antiviral or antifungal agents

Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;

Query Match 100.0%; Score 2785; DB 6; Length 5784;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAGGACTTTGACCGTGATGACCGGGTGATATTTGTAAAGATGATTTTGTGCAATGACAAC	60
Db	3000	GAGGACTTTGACCGTGATGACCGGGTGATATTTGTAAAGATGATTTTGTGCAATGACAAC	3059
Qy	61	ATCCGACATATTGATGATGTGTGCTTGAAAAAACAATGCCATCAGTGTGAGACAGACTTCAGG	120
Db	3060	ATCCGACATATTGATGATGTGTGCTTGAAAAAACAATGCCATCAGTGTGAGACAGACTTCAGG	3119
Qy	121	AACCTTCAGATGGTCCCTTGGATCCCAAGGAGCACCCAAAATTGATCCCAACTGGGTC	180
Db	3120	AACCTTCAGATGGTCCCTTGGATCCCAAGGAGCACCCAAAATTGATCCCAACTGGGTC	3179
Qy	181	ATTCCGCAATCAAGCGAAGGAGCTGGTTCAGACAGCCAACTCGGACCCCGGGCATGCTGTA	240
Db	3180	ATTCCGCAATCAAGCGAAGGAGCTGGTTCAGACAGCCAACTCGGACCCCGGGCATGCTGTA	3239
Qy	241	GGTTTTGACGAGTTTGGGTCTGTGGACCTTCAGTGGCACATTTCTACGTAAACACTGACCGG	300
Db	3240	GGTTTTGACGAGTTTGGGTCTGTGGACCTTCAGTGGCACATTTCTACGTAAACACTGACCGG	3299
Qy	301	GACGACGACTATGCTGGCTTCGTCTTTGGTTTACAGTCAAGCAGCCGCTTCTATGTGGTG	360
Db	3300	GACGACGACTATGCTGGCTTCGTCTTTGGTTTACAGTCAAGCAGCCGCTTCTATGTGGTG	3359
Qy	361	ATGTGGAGCAGGTGACGACAGACCTTACTGGGAGGACGAGCCACGCGGGGCTATGGTAC	420
Db	3360	ATGTGGAGCAGGTGACGACAGACCTTACTGGGAGGACGAGCCACGCGGGGCTATGGTAC	3419
Qy	421	TCCGGCGTGTCCCTCAAGTGGTGTAATCCACACAGGGGACGGCGGAGCACCCTGAGGAAC	480
Db	3420	TCCGGCGTGTCCCTCAAGTGGTGTAATCCACACAGGGGACGGCGGAGCACCCTGAGGAAC	3479
Qy	481	CGCTGTGGCACAGGGGAAACACCGCGGGCAGGTGGGAACTTTATGGCAGACCCGAGG	540
Db	3480	CGCTGTGGCACAGGGGAAACACCGCGGGCAGGTGGGAACTTTATGGCAGACCCGAGG	3539
Qy	541	AACATTGGCTGGAAGGACTACAGCGCCCTATAGGTGGCACCTTGACTCAGGCCCCAAGACC	600
Db	3540	AACATTGGCTGGAAGGACTACAGCGCCCTATAGGTGGCACCTTGACTCAGGCCCCAAGACC	3599
Qy	601	GGCTACATCAGAGCTTATAGTCGATGAAGGAAAAACAGGTCTATGGCAGACTCAGGACCTATC	660
Db	3600	GGCTACATCAGAGCTTATAGTCGATGAAGGAAAAACAGGTCTATGGCAGACTCAGGACCTATC	3659
Qy	661	TATGACCAAAACCTACGCTGGCGGGCGGTCTATTGTCTCTCTCAAGAAATGGTC	720
Db	3660	TATGACCAAAACCTACGCTGGCGGGCGGTCTATTGTCTCTCTCAAGAAATGGTC	3719
Qy	721	TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTGCTGCATTTCCG	780

Db 4800 TATTTTAAATATGCGCTAGTAACACATATGATGATATTAATTTCTAGAACAAACATCTAATA 4859
QY 1861 AGTATATAATCCCTGTCGAAATATGAGCTTGTGTAATATATAGGTTGTGTCAGCATGAACATG 1920
Db 4860 AGTATATAATCCCTGTCGAAATATGAGCTTGTGTAATATATAGGTTGTGTCAGCATGAACATG 4919
QY 1921 CTAGAAGCTGTACAGAAATACATAGAGAAATATGAGGAGTTTATGATGGAACCTTAATAT 1980
Db 4920 CTAGAAGCTGTACAGAAATACATAGAGAAATATGAGGAGTTTATGATGGAACCTTAATAT 4979
QY 1981 ATAATGTCGTCAGCGAATTTAGTTCATATTTGTTTCTGTTATCTGTTATCTGCTGTATATGG 2040
Db 4980 ATAATGTCGTCAGCGAATTTAGTTCATATTTGTTTCTGTTATCTGTTATCTGCTGTATATGG 5039
QY 2041 AATCTCTTTTAAATCAACGCTGAAACGAAATCAGCAATTTAGTCTGTCAGGAGCAACCCAA 2100
Db 5040 AATCTCTTTTAAATCAACGCTGAAACGAAATCAGCAATTTAGTCTGTCAGGAGCAACCCAA 5099
QY 2101 TAATCAGTCATGTTAAATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2160
Db 5100 TAATCAGTCATGTTAAATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 5159
QY 2161 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2220
Db 5160 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 5219
QY 2221 TAAGGGTTTGTAGTAAAGAGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2280
Db 5220 TAAGGGTTTGTAGTAAAGAGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5279
QY 2281 CCCTTTCTGTTAAATAGTCATCATCTTCTATGCCAAACAGGACGATCCATCACTTTAGT 2340
Db 5280 CCCTTTCTGTTAAATAGTCATCATCTTCTATGCCAAACAGGACGATCCATCACTTTAGT 5339
QY 2341 CTTAATGTACACATGCAATTTGTAATAAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTG 2400
Db 5340 CTTAATGTACACATGCAATTTGTAATAAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTG 5399
QY 2401 TTG 2460
Db 5400 TTG 5459
QY 2461 GAAGCGTTGGGATATCTCAATTAATGTAGCGACTGTCAACAGCGTGCAGGTTTCTGTTT 2520
Db 5460 GAAGCGTTGGGATATCTCAATTAATGTAGCGACTGTCAACAGCGTGCAGGTTTCTGTTT 5519
QY 2521 CTGTTCTGTTGGGTCACCGTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2580
Db 5520 CTGTTCTGTTGGGTCACCGTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5579
QY 2581 TGTACCAATATTTTGTAAATTTATTTATGTTTCTTAAACAAATTTATCGTATAGTTGA 2640
Db 5580 TGTACCAATATTTTGTAAATTTATTTATGTTTCTTAAACAAATTTATCGTATAGTTGA 5639
QY 2641 TGAACGTCATGTTGTTGTCGAAAGCTGTAAATTTATTTATGTTCTCATGTTCA 2700
Db 5640 TGAACGTCATGTTGTTGTCGAAAGCTGTAAATTTATTTATGTTCTCATGTTCA 5699
QY 2701 AATTTTCACTGTAACCCCTGACCTTAGCTAGACCTCATTTTAAAGATTAAACACAGG 2760
Db 5700 AATTTTCACTGTAACCCCTGACCTTAGCTAGACCTCATTTTAAAGATTAAACACAGG 5759
QY 2761 AAATAAATTTGTAACCAAGGTTTCT 2785
Db 5760 AAATAAATTTGTAACCAAGGTTTCT 5784

RESULT 3

ABK10277

ID ABK10277 standard; cDNA; 5784 BP.

XX

AC ABK10277;

XX

DT 21-MAY-2002 (first entry)
XX Human cDNA encoding thrombospondin 2 protein.
DE
XX
KW Human; Osteopontin 1; OPN1; ss; Gene; thrombospondin 2; wound response;
foreign body response; cut; abrasion; burn; vulnary.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 240..3758
FT /*tag= a
FT /product= "Thrombospondin 2 protein"
XX
XX WO200209735-A2.
PN
XX
XX 07-FEB-2002.
XX
XX 31-JUL-2001; 2001WO-US024147.
PF
XX
XX 01-AUG-2000; 2000US-0222071P.
PR
XX
XX (UNIW) UNIV WASHINGTON.
XX
XX Bornstein P, Kyriakides T, Ratner B, Giachelli C, Martinson L;
PI Scatena M;
PI
XX
XX WPI; 2002-217098/27.
DR P-PSDB; AAU76902.
XX
XX Modulating the amount or biological activity of thrombospondin 2 or
osteopontin in an animal for modulating a wound response, comprises
introducing osteopontin or thrombospondin 2 antagonist into an animal.
XX
XX Disclosure; Page 50-54; 54pp; English.
XX
XX This invention relates to a method for modulating the amount or
biological activity of thrombospondin 2 or osteopontin in an animal. The
method involves introducing into the animal an osteopontin or
thrombospondin 2 antagonist. Using the methods of the invention the
amount or biological activity of thrombospondin 2 or osteopontin protein
may be modulated. The method of the invention is useful for modulating
the amount or biological activity of thrombospondin 2 or osteopontin in
an animal which exhibits a wound response or a foreign body response,
where the method can be used to improve the wound response or reduces
the foreign body response. The method is useful for modulating a wound
response or for reducing a foreign body response in an animal and is also
useful for improving the wound response, such as at the site of a cut,
abrasion or burn. The present sequence represents the human cDNA encoding
the thrombospondin 2 protein used in the method of the invention
XX

SQ Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;

Query Match 100.0%; Score 2785; DB 6; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTGGACGGTGTGACGGGGTGATATTTGTAAGATGATTTTGACATCAAC 60
Db 3000 GAGGACTTGGACGGTGTGACGGGGTGATATTTGTAAGATGATTTTGACATCAAC 3059
QY 61 ATCCACATATTTGATGATGTTGTCTCTGTAACCAATGCCATCAGTGACAGATTGAGG 120
Db 3060 ATCCACATATTTGATGATGTTGTCTCTGTAACCAATGCCATCAGTGACAGATTGAGG 3119
QY 121 AACTTCCAGATGTTGCTTGGATCCCAAGGGACCAACCAATTTGATCCCACTGGGTC 180
Db 3120 AACTTCCAGATGTTGCTTGGATCCCAAGGGACCAACCAATTTGATCCCACTGGGTC 3179
QY 181 ATTCGCCATCAAGGCAAGGAGCTGGTTTCAGACAGCCAACTCGGACCCCGGATCGCTGTA 240
Db 3180 ATTCGCCATCAAGGCAAGGAGCTGGTTTCAGACAGCCAACTCGGACCCCGGATCGCTGTA 3239

QY	241	GGTTTTGACAGATTTGGGTCCTGFGAGCTTCAGTGGGACATTTCTAGTAAACACTGACCGG	300
DB	3240	GGTTTTGACAGATTTGGGTCCTGFGAGCTTCAGTGGGACATTTCTAGTAAACACTGACCGG	3299
QY	301	GACGACGACTATGCTGGCTTCGTCTTTGGTTACCACTCAAGCAGCGCTTCATGTGGTG	360
DB	3300	GACGACGACTATGCTGGCTTCGTCTTTGGTTACCACTCAAGCAGCGCTTCATGTGGTG	3359
QY	361	ATGTGGAAGCAGGTGACGACAGCTACTGGGAGGACAGGCCACCGGGCCCTATGGCTAC	420
DB	3360	ATGTGGAAGCAGGTGACGACAGCTACTGGGAGGACAGGCCACCGGGCCCTATGGCTAC	3419
QY	421	TCGGGCGTGCTCCCTCAAGFTGGTGAACCTCCAACAACGGGGACGGGCGAGCACCTGAGGAAC	480
DB	3420	TCGGGCGTGCTCCCTCAAGFTGGTGAACCTCCACACAACGGGGACGGGCGAGCACCTGAGGAAC	3479
QY	481	CGGCTGTGGCACACGGGGAAACACCGCGGGCAGGTGCGAACCTTATGGCAGCACCCACAGG	540
DB	3480	CGGCTGTGGCACACGGGGAAACACCGCGGGCAGGTGCGAACCTTATGGCAGCACCCACAGG	3539
QY	541	AACATTTGGCTGGAGGACTACAGGGCCTATAGGTGGCACCTGACTCACAGGCCCAAGACC	600
DB	3540	AACATTTGGCTGGAGGACTACAGGGCCTATAGGTGGCACCTGACTCACAGGCCCAAGACC	3599
QY	601	GGCTACATCAGAGTCTTAGTGCAATGAAGAAAAACAGGTCAATGGCAGACTCAGGACCTATC	660
DB	3600	GGCTACATCAGAGTCTTAGTGCAATGAAGAAAAACAGGTCAATGGCAGACTCAGGACCTATC	3659
QY	661	TATGACCAACCTACGCTGGGGGGCGCTGGGTCTATTGTCTTCTCAAGAAATGGTTC	720
DB	3660	TATGACCAACCTACGCTGGGGGGCGCTGGGTCTATTGTCTTCTCAAGAAATGGTTC	3719
QY	721	TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTGCTGCAATTCGG	780
DB	3720	TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTGCTGCAATTCGG	3779
QY	781	GCAATGGCCCTGTGATGCCAATGGTCCCTAGACACCTCAGTTCAATTTGGTTCCTGGGCT	840
DB	3780	GCAATGGCCCTGTGATGCCAATGGTCCCTAGACACCTCAGTTCAATTTGGTTCCTGGGCT	3839
QY	841	TCTCTCTCAGCAGCACTCTGTGTCCTTGACCTTAACCTGATGGTTCCTCACTCCCTCG	900
DB	3840	TCTCTCTCAGCAGCACTCTGTGTCCTTGACCTTAACCTGATGGTTCCTCACTCCCTCG	3899
QY	901	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATATAATCAATGGAACCTCAGAGATGAA	960
DB	3900	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATATAATCAATGGAACCTCAGAGATGAA	3959
QY	961	CATCTAACCCCACTAGAGAAACCAAGTTTGGTGATATATGAGACCTTTATGTGGAGTGA	1020
DB	3960	CATCTAACCCCACTAGAGAAACCAAGTTTGGTGATATATGAGACCTTTATGTGGAGTGA	4019
QY	1021	TTGGGCATGGCAATTAACATTTGCTTTTCTTTGTTTCTTTTAAAGAAATCAGCTTTACATATA	1080
DB	4020	TTGGGCATGGCAATTAACATTTGCTTTTCTTTGTTTCTTTTAAAGAAATCAGCTTTACATATA	4079
QY	1081	AAATGTGAATTAATCTTATTTGTTATGTTATGGAGTTGAAGGGAATACTGTGSCATAAG	1140
DB	4080	AAATGTGAATTAATCTTATTTGTTATGTTATGGAGTTGAAGGGAATACTGTGSCATAAG	4139
QY	1141	CCATTATGATAAATTAAGCATGAAAAATATTGCTGAACTACTTTTGGTGCCTTAAAGTTGT	1200
DB	4140	CCATTATGATAAATTAAGCATGAAAAATATTGCTGAACTACTTTTGGTGCCTTAAAGTTGT	4199
QY	1201	CACATTCTTGAATTAGAGTTGCTCTACAATGACACACAAATCCGCTAAATAAATATA	1260
DB	4200	CACATTCTTGAATTAGAGTTGCTCTACAATGACACACAAATCCGCTAAATAAATATA	4259
QY	1261	AACAAGGGTCAATTTCAAATTTGCAAGTAAATGTTTTAGTAAAGAGAGATTTAGAGACAACAG	1320
DB	4260	AACAAGGGTCAATTTCAAATTTGCAAGTAAATGTTTTAGTAAAGAGAGATTTAGAGACAACAG	4319
QY	1321	GCATGACCAATGACATPAAGCTACCGATTAACCTAATCGGAAACATGTAAACGCTTACAAAA	1380

Db	4320	GCATAGCAAAATGACATAAGAGCTACCGAATTAACCTATATCGGAAACATGTAAACAGTTACAAA	4379
QY	1381	ATAAACGAACCTCTCCTTCTTGTCCTACAATGAAAGCCCTCATGTGCAGTAGAGATCGAGTT	1440
Db	4380	ATAAACGAACCTCTCCTTCTTGTCCTACAATGAAAGCCCTCATGTGCAGTAGAGATCGAGTT	4439
QY	1441	TCATCAAGAAACAAACATCCTTCGCAATGGGTGTGACGGGTTCAGATGTGGATTTGCG	1500
Db	4440	TCATCAAGAAACAAACATCCTTCGCAATGGGTGTGACGGGTTCAGATGTGGATTTGCG	4499
QY	1501	AAAAACCTCATTTTAAGTAAAGAGTTAGCAGAGCAAGTGCGGTGTCTTTAGCTGCTGCTGT	1560
Db	4500	AAAAACCTCATTTTAAGTAAAGAGTTAGCAGAGCAAGTGCGGTGTCTTTAGCTGCTGCTGT	4559
QY	1561	GCCGTTGTGCGGTGCGGGAGGCTCCTGCGTGAAGTTCCTTCCCAGCTTTGCTGCGCTGAG	1620
Db	4560	GCCGTTGTGCGGTGCGGGAGGCTCCTGCGTGAAGTTCCTTCCCAGCTTTGCTGCGCTGAG	4619
QY	1621	AGGAACAGAGAGCAGACGACAGCGCGGAAAGCGCATCTAACCGCTATCTAGGCTTTTCG	1680
Db	4620	AGGAACAGAGAGCAGACGACAGCGCGGAAAGCGCATCTAACCGCTATCTAGGCTTTTCG	4679
QY	1681	TAACTGCGGACAAGTGTCTTTTACCTGATTTGATGATACATTTCAATAAGGTTCCAGTTA	1740
Db	4680	TAACTGCGGACAAGTGTCTTTTACCTGATTTGATGATACATTTCAATAAGGTTCCAGTTA	4739
QY	1741	TAAATATTTTGTGTTAAATATTTTAAAGTGACTATAGAAATGCACTCCATTTACCAAGTAAC	1800
Db	4740	TAAATATTTTGTGTTAAATATTTTAAAGTGACTATAGAAATGCACTCCATTTACCAAGTAAC	4799
QY	1801	TATTTTAAATATGCTTAGTAAACACATATGTAGTATATTAATTTCTAGAAACAAACATCTAATA	1860
Db	4800	TATTTTAAATATGCTTAGTAAACACATATGTAGTATATTAATTTCTAGAAACAAACATCTAATA	4859
QY	1861	AGTATATAATCCTGTGAAAATATGAGGCTTGATATAATATTAGGTTGTGCAGATGAAGCATG	1920
Db	4860	AGTATATAATCCTGTGAAAATATGAGGCTTGATATAATATTAGGTTGTGCAGATGAAGCATG	4919
QY	1921	CTAGAGCTGTAAACAGAAATACATAGAGAAATAATGAGGAGTTTATGATGGAACTTAATAT	1980
Db	4920	CTAGAGCTGTAAACAGAAATACATAGAGAAATAATGAGGAGTTTATGATGGAACTTAATAT	4979
QY	1981	ATAATGTTGCCAGCGATTTTAGTTTCAATATTTGTTACTGTATCTAATCTGCTGTATATGG	2040
Db	4980	ATAATGTTGCCAGCGATTTTAGTTTCAATATTTGTTACTGTATCTAATCTGCTGTATATGG	5039
QY	2041	AATCTCTTTTAATTCMAAGCTGAAAACGAAATCAGCATTTAGTCTTGCCAGGACACCCAA	2100
Db	5040	AATCTCTTTTAATTCMAAGCTGAAAACGAAATCAGCATTTAGTCTTGCCAGGACACCCAA	5099
QY	2101	TAAATCAGTCAATGTGTAATATATGCAAAAGTTGTTTTGTTTTTTTTTTGTTGTTGTTGG	2160
Db	5100	TAAATCAGTCAATGTGTAATATATGCAAAAGTTGTTTTGTTTTTTTTTTGTTGTTGTTGG	5159
QY	2161	TTTTTTTGTGTTTAAGTTCGATGATCTTTCGACAGAAATAGTCACTCATCCCATCCACA	2220
Db	5160	TTTTTTTGTGTTTAAGTTCGATGATCTTTCGACAGAAATAGTCACTCATCCCATCCACA	5219
QY	2221	TAAAGGGTTTAGTAAAGAGAGTCTGTCTGTCTGATGGATAGGGGGCAAAATCTTTTTC	2280
Db	5220	TAAAGGGTTTAGTAAAGAGAGTCTGTCTGTCTGATGGATAGGGGGCAAAATCTTTTTC	5279
QY	2281	CCCTTCTGTGTTAAATAGTCATCAATTTCTATGCCAAACAGAAACGATCCATAACTTTAGT	2340
Db	5280	CCCTTCTGTGTTAAATAGTCATCAATTTCTATGCCAAACAGAAACGATCCATAACTTTAGT	5339
QY	2341	CTTAATGTACAAATTTGCAATTTTATGATAAAATTAATTTTGTCTGTTTCCCTTTGAGGTTGATCG	2400
Db	5340	CTTAATGTACAAATTTGCAATTTTATGATAAAATTAATTTTGTCTGTTTCCCTTTGAGGTTGATCG	5399
QY	2401	TTGTGTTGTTTGTGCTGCACTTTTATCTTTTGTGCGTGTGGAGCTGTATTTCCCGAGACAAC	2460

Db	3900	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAAATATCAATCGAACTCAGAGATGAA	3959
Qy	961	CATCTAACCCCATAGAGAGAAACCAAGTTGGTGATATATGAGACATTTATGTGAGTGAATA	1020
Db	3960	CATCTAACCCCATAGAGAGAAACCAAGTTGGTGATATATGAGACATTTATGTGAGTGAATA	4019
Qy	1021	TTGGGCATGCCATTACATTCGCTTTTCTGTGTTTAAAAAGAAATGACGTTTACATATA	1080
Db	4020	TTGGGCATGCCATTACATTCGCTTTTCTGTGTTTAAAAAGAAATGACGTTTACATATA	4079
Qy	1081	AAATGTAATTTACTTTATTTGTTATTTATGAGTTTGAAGGGAATACTGTGCATAAG	1140
Db	4080	AAATGTAATTTACTTTATTTGTTATTTATGAGTTTGAAGGGAATACTGTGCATAAG	4139
Qy	1141	CCATTTATGATAAATTAAGCATGAAAAATATTGCTGAACTACTCTTTGGTGCTTTAAAGTTGT	1200
Db	4140	CCATTTATGATAAATTAAGCATGAAAAATATTGCTGAACTACTCTTTGGTGCTTTAAAGTTGT	4199
Qy	1201	CACATTTCTTGAATTTAGAGTTGCTCTPACAATGACACACAAATCCCGCTAAAAATTATA	1260
Db	4200	CACATTTCTTGAATTTAGAGTTGCTCTPACAATGACACACAAATCCCGCTAAAAATTATA	4259
Qy	1261	AACAAGGGTCAATTTCAAAATTTGAAGTAATGTTTTAGTAAGGAGAGATTAGAACACAACAG	1320
Db	4260	AACAAGGGTCAATTTCAAAATTTGAAGTAATGTTTTAGTAAGGAGAGATTAGAACACAACAG	4319
Qy	1321	GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAACACAGTTACAAA	1380
Db	4320	GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAACACAGTTACAAA	4379
Qy	1381	ATAAACGAATCTCTCTCTCTCTCTCAATGAAAGCCCTCATGTGCAAGTACAGATGCAGTT	1440
Db	4380	ATAAACGAATCTCTCTCTCTCTCTCAATGAAAGCCCTCATGTGCAAGTACAGATGCAGTT	4439
Qy	1441	TCATCAAGAACAAAACATCTCTGCAATGGGTGTGACGCGGTTCCAGATGTGGAATTGGC	1500
Db	4440	TCATCAAGAACAAAACATCTCTGCAATGGGTGTGACGCGGTTCCAGATGTGGAATTGGC	4499
Qy	1501	AAAAACCTCATTTAAGTAAAAAGTTAGCAGAGCAAAAGTGCGBGTCTTTAGCTGCTGTTGT	1560
Db	4500	AAAAACCTCATTTAAGTAAAAAGTTAGCAGAGCAAAAGTGCGBGTCTTTAGCTGCTGTTGT	4559
Qy	1561	GCCGTTGTGGGTGCGGAGGCTCCTGCCTGAGCTCTTCCCAGCTTTTGTGCTGCTGAG	1620
Db	4560	GCCGTTGTGGGTGCGGAGGCTCCTGCCTGAGCTCTTCCCAGCTTTTGTGCTGCTGAG	4619
Qy	1621	AGGAACAGAGCAGACGCAAGGCGGAAAGGCGCATCTAAACGCGTATCTAGGCTTTGG	1680
Db	4620	AGGAACAGAGCAGACGCAAGGCGGAAAGGCGCATCTAAACGCGTATCTAGGCTTTGG	4679
Qy	1681	TAACTCGGCAACAGTTGCTTTTAACTGATTGATGATACATTTCAATTAAGGTTCCAGTTA	1740
Db	4680	TAACTCGGCAACAGTTGCTTTTAACTGATTGATGATACATTTCAATTAAGGTTCCAGTTA	4739
Qy	1741	TAAATATTTTCTTAAATTTTAAAGTGACTATAGAAATGCAATCCCAATTTACCAAGTAACT	1800
Db	4740	TAAATATTTTCTTAAATTTTAAAGTGACTATAGAAATGCAATCCCAATTTACCAAGTAACT	4799
Qy	1801	TATTTTAAATATGCTAGTAAACATATGTAGTATAATTTCTAGAAACAAACATCTAATA	1860
Db	4800	TATTTTAAATATGCTAGTAAACATATGTAGTATAATTTCTAGAAACAAACATCTAATA	4859
Qy	1861	AGTATATTAATCTGTGAAAAATAGAGGCTTGATTAATTTAGGTTGTGACGATGAGCATG	1920
Db	4860	AGTATATTAATCTGTGAAAAATAGAGGCTTGATTAATTTAGGTTGTGACGATGAGCATG	4919
Qy	1921	CTAGAGCTGTAAACAGATACATAGAGAAATAGAGGAGTTTATGATGGAAACCTTAATAT	1980
Db	4920	CTAGAGCTGTGTAAACAGATACATAGAGAAATAGAGGAGTTTATGATGGAAACCTTAATAT	4979
Qy	1981	ATAATGTTGCCAGGCAATTTTGTGTTCAATATTTTGTACTGTATCTATCTGCTGTATAGG	2040

Db	4980	ATAATGTTGCCAGCGATTTTAGTTCAAAATTTGTGTACTGTATCTATCTGCTGTATATGG	5033
Qy	2041	AAATCTTTTAAATTCAAACCGTGA AAAACGAATCAGCAATTTAGTCTTTGCCAGGCACCCCAA	2100
Db	5040	AAATCTTTTAAATTCAAACCGTGA AAAACGAATCAGCAATTTAGTCTTTGCCAGGCACCCCAA	5099
Qy	2101	TAAATCAGTCATGCTGAATATGCAACAAGTTTGTGTTTTTGTGTTTTTTTCTGCTGGTGG	2160
Db	5100	TAAATCAGTCATGCTGAATATGCAACAAGTTTGTGTTTTTGTGTTTTTTTGTGCTGGTGG	5159
Qy	2161	TTTTTTTTGCTTTAAAGTTTGCAATGATCTTTCTGCAGGAAATAGTCACTCATCCCACTCCACA	2220
Db	5160	TTTTTTTTGCTTTAAAGTTTGCAATGATCTTTCTGCAGGAAATAGTCACTCATCCCACTCCACA	5219
Qy	2221	TAAGGGGTTTGTAGTAAGAGAAAGTCGTCTGTCTGATGATGGATAGGGGGGAAATCTTTTTTC	2280
Db	5220	TAAGGGGTTTGTAGTAAGAGAAAGTCGTCTGTCTGATGATGGATAGGGGGGAAATCTTTTTTC	5279
Qy	2281	CCCTTTCTGTTAATAGTCATCACATTTCTATGTCCTATGCGAAGCAAGATCATAACTTTAGT	2340
Db	5280	CCCTTTCTGTTAATAGTCATCACATTTCTATGTCCTATGCGAAGCAAGATCATAACTTTAGT	5339
Qy	2341	CTTAATGTCACACATGTCATTTTGCATAAAATTAATTTTGTGTTTCTCTTTGAGGTTGATCG	2400
Db	5340	CTTAATGTCACACATGTCATTTTGCATAAAATTAATTTTGTGTTTCTCTTTGAGGTTGATCG	5399
Qy	2401	TTGRTGTTGTTTGTGTCACATTTTACATTTTTCGCTGTGGAGCTGTATTTCCCGAGACAAC	2460
Db	5400	TTGRTGTTGTTTGTGTCACATTTTACATTTTTCGCTGTGGAGCTGTATTTCCCGAGACAAC	5459
Qy	2461	GAAAGGTTGGGATCTTCAATTAATGTAGCGACTGTCAACAGCGTGCAGGTTTTCGTGTT	2520
Db	5460	GAAAGGTTGGGATCTTCAATTAATGTAGCGACTGTCAACAGCGTGCAGGTTTTCGTGTT	5519
Qy	2521	CTGTGTTGTGGGGTCAACCGGTACAAATGGTGTGGGAATGACGATGTGAATATTTAGAA	2580
Db	5520	CTGTGTTGTGGGGTCAACCGGTACAAATGGTGTGGGAATGACGATGTGAATATTTAGAA	5579
Qy	2581	TGTACCAATATTTTTGTGTAATAATTTATGTTTTTTCTAAACAAATTTATCGTATAGTTGA	2640
Db	5580	TGTACCAATATTTTTGTGTAATAATTTATGTTTTTTCTAAACAAATTTATCGTATAGTTGA	5639
Qy	2641	TGAAACGTCATGCTGTTTTTGGCAAGACTGTAAATAATTTATTTATGTTTCCACATGGTCAA	2700
Db	5640	TGAAACGTCATGCTGTTTTTGGCAAGACTGTAAATAATTTATTTATGTTTCCACATGGTCAA	5699
Qy	2701	AAATTTCAACCACTGA AACCCCTGCACCTAGCTAGTA GAACCTCACTTTTAAAGATTAA CAACAGG	2760
Db	5700	AAATTTCAACCACTGA AACCCCTGCACCTAGCTAGTA GAACCTCACTTTTAAAGATTAA CAACAGG	5759
Qy	2761	AAATAAATGTAAAAAGGTTTTTCT	2785
Db	5760	AAATAAATGTAAAAAGGTTTTTCT	5784

PRESENT 5

RESULTS 3
ADD40611

AAD406LI
IR 33D40C11 standard. DNA. 5784 BP

ID AAD40611

XX

AC AAD40611;

XX

DT 30-OCT-2002 (first entry)

XX
T

Human thrombospondin (TSP)-2 DNA.

DE
HUMAN
CYTOMEGALOVIRUS (HCMV) - 2 ENNA
v v

XX

KW Human; cardiovascular disease; CV

KW anticoagulant therapy; atherosclerosis

KW myocardial infarction; venous thr

KW peripheral vascular disease; cong

Key	Location/Qualifiers
FT CDS	240..3758
FT	/*tag= a
FT	/product= "TSP-2 protein"
FT variation	/replace(3949, G)
FT	/*tag= b
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
XX	
PN WO200239122-A2.	
XX	
PD 16-MAY-2002.	
XX	
PP 09-NOV-2001; 2001WO-US047406.	
XX	
PR 09-NOV-2000; 2000US-0248185P.	
PR 22-DEC-2000; 2000US-0257417P.	
XX	
PA (MILL-) MILLENNIUM PHARM INC.	
XX	
PI McCarthy J;	
XX	
DR WPI; 2002-519324/55.	
XX P-PSDB; RAE25031.	
XX	
PT Diagnosing and monitoring a cardiovascular disease in a patient, by	
PT assessing the level of thrombospondin protein and/or gene expression in a	
PT biological sample.	
XX	
XX Example 1; Page 100-102; 108pp; English.	
XX	
CC The invention relates to a method of diagnosing and monitoring a	
CC cardiovascular disease (CVD) in a patient. The method involves assessing	
CC the level of thrombospondin (TSP) protein and/or gene expression in a	
CC biological sample. The invention also provides a method useful for	
CC assessing the efficacy of a compound or anticoagulant therapy for	
CC inhibiting CVD in a patient. CVD include atherosclerosis, coronary artery	
CC disease, myocardial infarction (MI), stroke, peripheral vascular disease,	
CC congestive heart failure, venous thromboembolism and pulmonary embolism.	
CC The invention is also used in gene therapy. The present sequence is human	
CC TSP-2 DNA	
XX	
SQ Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;	
Query Match	100.0%; Score 2785; DB 6; Length 5784;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2785; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTGACGGTGATGACGGGGGTGATATTGTAAAGATGATTTTGACATGACAAAC 60	
DB 3000 GAGGACTTGACGGTGATGACGGGGGTGATATTGTAAAGATGATTTTGACATGACAAAC 3059	
QY 61 ATCCGAGATTCATGATGTGTCTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 120	
DB 3060 ATCCGAGATTCATGATGTGTCTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 3119	
QY 121 AACTTCCAGATGTGCCCTTGGATCCCAAGGGACCAACCCAAATTGATCCCACTGGGTC 180	
DB 3120 AACTTCCAGATGTGCCCTTGGATCCCAAGGGACCAACCCAAATTGATCCCACTGGGTC 3179	
QY 181 ATTGCGCATCAAGGCAAGGAGCTGGTTTCAGACAGCAACTCGGACCCCGGCATCGCTGTA 240	
DB 3180 ATTGCGCATCAAGGCAAGGAGCTGGTTTCAGACAGCAACTCGGACCCCGGCATCGCTGTA 3239	
QY 241 GGTTTTGACGAGTTTGGGTCCTGTGGACTTCAGTGGGCACATCTTACGTAAACACTGCCG 300	
DB 3240 GGTTTTGACGAGTTTGGGTCCTGTGGACTTCAGTGGGCACATCTTACGTAAACACTGCCG 3299	
QY 301 GAGCAGCATATCTGGCTTCGTCCTTTGGTTACAGTCCAGCAGCGCTTCTATGTGGTG 360	
DB 3300 GAGCAGCATATCTGGCTTCGTCCTTTGGTTACAGTCCAGCAGCGCTTCTATGTGGTG 3359	
QY 361 ATGTGGAACAGGTGACGAGACCTACTGGGAGGACAGCCCAACGCGGGGCTATGGCTAC 420	

Db	3240	TGACGAGTTTGGGTCCTGTGGACTTCAGTGGCCACATTTCTACGTAAACACTGACCGG	3299
Qy	3301	GAGCAGCACTATGCTGGCTTCGTCTTTTGGTTTACCAGTCAAGCAGCCGCTCTCTATGTGCGTG	360
Db	3300	GACGACCACTATGCTGGCTTCGTCTTTTGGTTTACCAGTCAAGCAGCCGCTCTCTATGTGCGTG	3359
Qy	361	ATGTGGGAAGCAGGTGACGCGACACTACTGGGAGGACCAGGCCACCGGGGCTATGGCTAC	420
Db	3360	ATGTGGGAAGCAGGTGACGCGACACTACTGGGAGGACCAGGCCACCGGGGCTATGGCTAC	3419
Qy	421	TCGGCGGTGTCCTCAAGTGTGTAACTCCACACCGGGGACGGGCGAGCACTTGAGGAAC	480
Db	3420	TCGGCGGTGTCCTCAAGTGTGTAACTCCACACCGGGGACGGGCGAGCACTTGAGGAAC	3479
Qy	481	GGCTGTGTGGCACACGGGGAAACACGCCGGGGCAGGTGGGAACCTTATGCGCAGCACCCGAG	540
Db	3480	GGCTGTGTGGCACACGGGGAAACACGCCGGGGCAGGTGGGAACCTTATGCGCAGCACCCGAG	3539
Qy	541	AACATTGGCTGGAAGGACTACAGCGGCTATAGTGGCACCTGACTCACAGGCCCAAGACC	600
Db	3540	AACATTGGCTGGAAGGACTACAGCGGCTATAGTGGCACCTGACTCACAGGCCCAAGACC	3599
Qy	601	GGCTACATCAGAGTCTTAGTGCATGAGGAAACAGGTCAATGCGCAGACTCAGAACCTATC	660
Db	3600	GGCTACATCAGAGTCTTAGTGCATGAGGAAACAGGTCAATGCGCAGACTCAGAACCTATC	3659
Qy	661	TATGACCAACCTACGCTGGCGGGCGGCTGGGTCTATTGTCTTCTCAAGAAATGGTTC	720
Db	3660	TATGACCAACCTACGCTGGCGGGCGGCTGGGTCTATTGTCTTCTCAAGAAATGGTTC	3719
Qy	721	TATTTCTCAGACCTCAAGTACGAATGCAAGTATTTAAACAAGATTGTGCTCATTTCCG	780
Db	3720	TATTTCTCAGACCTCAAGTACGAATGCAAGTATTTAAACAAGATTGTGCTCATTTCCG	3779
Qy	781	GCAATGCCCTGTGCATGCCCATGTGCTCCTAGACACCTCAGTTGATGTGCTCTTGGCGGT	840
Db	3780	GCAATGCCCTGTGCATGCCCATGTGCTCCTAGACACCTCAGTTGATGTGCTCTTGGCGGT	3839
Qy	841	TCCTCTCTAGCAGCACCTCCTGTCCCTTGACCTTAACCTCTGATGTGTTCTTCACCTCCCTG	900
Db	3840	TCCTCTCTAGCAGCACCTCCTGTCCCTTGACCTTAACCTCTGATGTGTTCTTCACCTCCCTG	3899
Qy	901	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAATCAGAGATGAA	960
Db	3900	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAATCAGAGATGAA	3959
Qy	961	CATCTAACCCACTAGAGGAAACAGTTTGGTGATATAGACACTTATGTGGAGTGAAAA	1020
Db	3960	CATCTAACCCACTAGAGGAAACAGTTTGGTGATATAGACACTTATGTGGAGTGAAAA	4019
Qy	1021	TTGGGCATGCCATTAATTTGCTTTTCTTTGTTTAAAAAGAAATGACGTTTACATATA	1080
Db	4020	TTGGGCATGCCATTAATTTGCTTTTCTTTGTTTAAAAAGAAATGACGTTTACATATA	4079
Qy	1081	AAATGTAAATTAATTTATTTATTTATGTGATATGAGTTTGAAGGAAATCTGTGCATAAG	1140
Db	4080	AAATGTAAATTTACTTATTTGATTTATGTGATATGAGTTTGAAGGAAATCTGTGCATAAG	4139
Qy	1141	CCATTATGATAAATTAAGCATCAAAAATTTGCTGAACTACTTTTGGTGTCTTAAAGTTGT	1200
Db	4140	CCATTATGATAAATTAAGCATCAAAAATTTGCTGAACTACTTTTGGTGTCTTAAAGTTGT	4199
Qy	1201	CACATTTCTTGAATTTAGAGTTGCTCTACAAATGACACACAAATCCCGCTTAAATTAATA	1260
Db	4200	CACATTTCTTGAATTTAGAGTTGCTCTACAAATGACACACAAATCCCGCTTAAATTAATA	4259
Qy	1261	AACAAGGGTCAATTTCAAATTTGAAATGTAATTTTATAGTAAGGAGAGATTAGAGACAAAG	1320
Db	4260	AACAAGGGTCAATTTCAAATTTGAAATGTAATTTTATAGTAAGGAGAGATTAGAGACAAAG	4319
Qy	1321	GCATAGCAAAATCACATAGCTACCGATTAACTAATCGGAAACATGTAAACAGTTTCAAAA	1380

Db	4320	GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAAGTTTACAAAA	4379
Qy	1381	ATAAAGCAACTCTCTCTTGTCTTACAAATGAAAGCCCTCATGTGCAGTAGAGATGCAAGTT	1440
Db	4380	ATAAAGCAACTCTCTCTTGTCTTACAAATGAAAGCCCTCATGTGCAGTAGAGATGCAAGTT	4439
Qy	1441	TCATCAAAAGAAACAAACATCCTTTGCAAATGGGTGTGAGCGGTTCCAGATGTGGAATTTGGC	1500
Db	4440	TCATCAAAAGAAACAAACATCCTTTGCAAATGGGTGTGAGCGGTTCCAGATGTGGAATTTGGC	4499
Qy	1501	AAAACCTCATTTAAAGTAAAGGTTAGCAGAGCAAAAGTCGCGTGCTTTAGCTGCTGCTTGT	1560
Db	4500	AAAACCTCATTTAAAGTAAAGGTTAGCAGAGCAAAAGTCGCGTGCTTTAGCTGCTGCTTGT	4559
Qy	1561	GCCGTTGTGGCGTGGGGAGGCTCTGCCTGAGCTTCCTTCCCAGCTTTTGTCTGCCCTGAG	1620
Db	4560	GCCGTTGTGGCGTGGGGAGGCTCTGCCTGAGCTTCCTTCCCAGCTTTTGTCTGCCCTGAG	4619
Qy	1621	AGGAACACAGACAGACAGCGACAGGCGCGAAAGGCGCATCTTAACGCGTATCTAGGCTTTGG	1680
Db	4620	AGGAACACAGACAGACAGCGACAGGCGCGAAAGGCGCATCTTAACGCGTATCTAGGCTTTGG	4679
Qy	1681	TAACTGGGACAAGTTGCTTTTACCTGATTGATGATACATTTTCATTAAGGTTTCCAGTTA	1740
Db	4680	TAACTGGGACAAGTTGCTTTTACCTGATTGATGATACATTTTCATTAAGGTTTCCAGTTA	4739
Qy	1741	TAAATATTTTGTTTAAATATTTTAAAGTGACTATAGAAATGCAATCCCATTTACAGTAACCT	1800
Db	4740	TAAATATTTTGTTTAAATATTTTAAAGTGACTATAGAAATGCAATCCCATTTACAGTAACCT	4799
Qy	1801	TATTTTAAATATGCTAGTAACACATATGTAGTATATTTCTAGAAACAAACATCTCAATA	1860
Db	4800	TATTTTAAATATGCTAGTAACACATATGTAGTATATTTCTAGAAACAAACATCTCAATA	4859
Qy	1861	AGTATATAATCCTGTGAAAAATATGAGCTTTGATAATATTAGTGTGTGACGATGAACATG	1920
Db	4860	AGTATATAATCCTGTGAAAAATATGAGCTTTGATAATATTAGTGTGTGACGATGAACATG	4919
Qy	1921	CTAGAAGCTGTACAGAAATACATAGAGATAATATGAGGAGTTTAATGATGAAACCTTAATAT	1980
Db	4920	CTAGAAGCTGTACAGAAATACATAGAGATAATATGAGGAGTTTAATGATGAAACCTTAATAT	4979
Qy	1981	ATAATGTTGCCAGGATTTTATGTTTCAATATTTGTTACTGTATCTATCTGCTGTATATGG	2040
Db	4980	ATAATGTTGCCAGGATTTTATGTTTCAATATTTGTTACTGTATCTATCTGCTGTATATGG	5039
Qy	2041	AATCTTTTAAATCAACCGCTGAAAAAGCAATCAGCAATTTAGTCTTGCCAGGCACACCCAA	2100
Db	5040	AATCTTTTAAATCAACCGCTGAAAAAGCAATCAGCAATTTAGTCTTGCCAGGCACACCCAA	5099
Qy	2101	TAAATCAGTCATGTATAATATGACAAAGTTTGTGTTTTTGTGTTTTTGTGTTTTTGTG	2160
Db	5100	TAAATCAGTCATGTATAATATGACAAAGTTTGTGTTTTTGTGTTTTTGTGTTTTTGTG	5159
Qy	2161	TTTTTTTTGCTTTAAGTTGCAATCTTTCTGAGGAAATAGTCACTCATCCCACTCCACA	2220
Db	5160	TTTTTTTTGCTTTAAGTTGCAATCTTTCTGAGGAAATAGTCACTCATCCCACTCCACA	5219
Qy	2221	TAAAGGGTTTATGTAAGAGAGTCTGTCTGTCTGATGATGGAATAGGGGGCAATCTTTTTTC	2280
Db	5220	TAAAGGGTTTATGTAAGAGAGTCTGTCTGTCTGATGATGGAATAGGGGGCAATCTTTTTTC	5279
Qy	2281	CCCTTTTCTGTTAATAGTCATCACTTTCTATGCCAAAACAGGAACGATCCATAACTTTAGT	2340
Db	5280	CCCTTTTCTGTTAATAGTCATCACTTTCTATGCCAAAACAGGAACGATCCATAACTTTAGT	5339
Qy	2341	CTTAATGTACAACTTCATTTTGTATAAAATTAATTTTGTCTTCTTCTTGTAGGTTGATCG	2400
Db	5340	CTTAATGTACAACTTCATTTTGTATAAAATTAATTTTGTCTTCTTCTTGTAGGTTGATCG	5399
Qy	2401	TTGTGTTGTTTTGCTGCACCTTTTACTTTTTTGTGCGTGTGGAGCTGTATTTCCGAGACAAC	2460
Db	5400	TTGTGTTGTTTTGCTGCACCTTTTACTTTTTTGTGCGTGTGGAGCTGTATTTCCGAGACAAC	5459

QY 2461 GAAGCGTTGGGATACCTTCATTAATGATGAGCACTGTCAACAGCGGTGCAGGTTTCTGTTT 2520
Db |||||||
QY 5460 GAAGCGTTGGGATACCTTCATTAATGATGAGCACTGTCAACAGCGGTGCAGGTTTCTGTTT 5519
Db |||||||
QY 2521 CTGTGTTGTGGGTCAACCGTACAAATGCTGGGAATGACGATGATGGAATATTAGAA 2580
Db |||||||
QY 5520 CTGTGTTGTGGGTCAACCGTACAAATGCTGGGAATGACGATGATGGAATATTAGAA 5579
Db |||||||
QY 2581 TGTACCAATATTTTGTAAATATTATGTTTCTTAAACAAATTTATCCGTATAGGTTGA 2640
Db |||||||
QY 5580 TGTACCAATATTTTGTAAATATTATGTTTCTTAAACAAATTTATCCGTATAGGTTGA 5639
Db |||||||
QY 2641 TGAACGTCATGTTTGGCAAGACTGTAAATATTATTTATGTTTCAATGTCACATGGTCAA 2700
Db |||||||
QY 5640 TGAACGTCATGTTTGGCAAGACTGTAAATATTATTTATGTTTCAATGTCACATGGTCAA 5699
Db |||||||
QY 2701 AATTTACCACTGAACCCCTGCACCTAGCTAGAGAACCTCATTTTAAAGATTAAACACAGG 2760
Db |||||||
QY 5700 AATTTACCACTGAACCCCTGCACCTAGCTAGAGAACCTCATTTTAAAGATTAAACACAGG 5759
Db |||||||
QY 2761 AATATAATTTGTAAGGTTTCT 2785
Db |||||||
QY 5760 AATATAATTTGTAAGGTTTCT 5784
Db |||||||

RESULT 8
AADS5115
ID AADS5115 standard; DNA; 5784 BP.
XX
AC AADS5115;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human THBS2 reference DNA (GI 307505).
XX
KW Thrombospondin 2; THBS2; angiotensin converting enzyme; polymorphism;
KW ACE-1; beta-fibrinogen; FGB; peripheral vascular disease; ischaemia;
KW vascular disease; myocardial infarction; pulmonary embolism; stroke;
KW atherosclerosis; coronary artery disease; venous thromboembolism; human;
KW gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 240..3758
FT /*tag= a
FT /product= "Human THBS2 reference protein"
FT replace(3949,G)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
XX WO2003020118-A2.
XX
XX PN
XX PD 13-MAR-2003.
XX
XX PF 04-SEP-2002; 2002WO-US028113.
XX
XX PR 05-SEP-2001; 2001US-0317178P.
XX PR 16-OCT-2001; 2001US-0329958P.
XX PR 14-DEC-2001; 2001US-00017724.
XX
XX PA (VITI-) VITIIVITY INC.
XX
XX PI Mccarthy J;
XX
XX DR WPI; 2003-300816/29.
XX DR P-PSDB; AAE36411.
XX
XX PT Identifying polymorphisms in thrombospondin 2, angiotensin converting
XX enzyme and/or beta-fibrinogen genes in nucleic acid sample of subject, by
XX contacting the nucleic acid with a complementary probe or primer.
XX

Claim 2; Fig 1; 194pp; English.

The invention relates to a method for determining the identity of one or more allelic variants of a polymorphic region of a thrombospondin 2 (THBS2), angiotensin converting enzyme (ACE)-1 and/or beta-fibrinogen (FGB) genes in a nucleic acid obtained from a subject. The method involves contacting the nucleic acid with a complementary probe or primer. The method is useful for diagnosing or aiding in the diagnosis of coronary artery disease, atherosclerosis, ischaemia, stroke, peripheral vascular disease, venous thromboembolism and pulmonary embolism. The present sequence is human THBS2 reference DNA

Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;

Query Match 100.0%; Score 2785; DB 9; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTGGACGGTGTGACGGGGTATATTGTTAAAGATGATTGACAATGACAAC 60
Db |||||||
QY 3000 GAGGACTTGGACGGTGTGACGGGGTATATTGTTAAAGATGATTGACAATGACAAC 3059
Db |||||||
QY 61 ATCCAGATATTGATGATGTGTCTCGAAACAAATGCCATCAGTGAGACAGACTTCAGG 120
Db |||||||
QY 3060 ATCCAGATATTGATGATGTGTCTCGAAACAAATGCCATCAGTGAGACAGACTTCAGG 3119
Db |||||||
QY 121 AACTTCCAGATGGTCCCTTTGGATCCCAAGGGACACCCAAATGTATCCCACTGGGTC 180
Db |||||||
QY 3120 AACTTCCAGATGGTCCCTTTGGATCCCAAGGGACACCCAAATGTATCCCACTGGGTC 3179
Db |||||||
QY 181 ATTGCCATCAAGCAAGGAGCTGTTTCAGACAGCACTCGGACCCCGGATCGCTGTA 240
Db |||||||
QY 3180 ATTGCCATCAAGCAAGGAGCTGTTTCAGACAGCACTCGGACCCCGGATCGCTGTA 3239
Db |||||||
QY 241 GGTTTGACGAGTTTGGGTCTGTGACTTCAGTGGGCACTTACGTAAACACATGACCGG 300
Db |||||||
QY 3240 GGTTTGACGAGTTTGGGTCTGTGACTTCAGTGGGCACTTACGTAAACACATGACCGG 3299
Db |||||||
QY 301 GACGACGACTATGCTGGCTTCGTCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGTG 360
Db |||||||
QY 3300 GACGACGACTATGCTGGCTTCGTCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGTG 3359
Db |||||||
QY 361 ATGTGGAAGAGGTGACGACACCTACTGCGAGGACACCCACGCGGCGCTTATGGCTAC 420
Db |||||||
QY 3360 ATGTGGAAGAGGTGACGACACCTACTGCGAGGACACCCACGCGGCGCTTATGGCTAC 3419
Db |||||||
QY 421 TCCGCGGTGTCCTCAAGTGTGTAACTCCACACGCGGACGCGGCGAGCACCTGAGGAAC 480
Db |||||||
QY 3420 TCCGCGGTGTCCTCAAGTGTGTAACTCCACACGCGGACGCGGCGAGCACCTGAGGAAC 3479
Db |||||||
QY 481 GCGCTGTGGCAACGCGGGAACACGCGCGGGCAGGTGGGAACTTATGGACGACCCCGAG 540
Db |||||||
QY 3480 GCGCTGTGGCAACGCGGGAACACGCGCGGGCAGGTGGGAACTTATGGACGACCCCGAG 3539
Db |||||||
QY 541 AACATTTGGTGAAGACTTACACGCGCTTATAGTGGCACTGACCTCACAGGCGCCAGACCC 600
Db |||||||
QY 3540 AACATTTGGTGAAGACTTACACGCGCTTATAGTGGCACTGACCTCACAGGCGCCAGACCC 3599
Db |||||||
QY 601 GGCTACATCAGAGTCTTATGTCATGAAGGAAACAGGTCATGGCAGACTCAGGACCTATC 660
Db |||||||
QY 3600 GGCTACATCAGAGTCTTATGTCATGAAGGAAACAGGTCATGGCAGACTCAGGACCTATC 3659
Db |||||||
QY 661 TATGACCAAACTTACGCTGGCGGGCGGTGGGTCTATTGTTCTTCTCAAGAAATGGTC 720
Db |||||||
QY 3660 TATGACCAAACTTACGCTGGCGGGCGGTGGGTCTATTGTTCTTCTCAAGAAATGGTC 3719
Db |||||||
QY 721 TATTTCTCAGACTCAAGTACCAATGTCAGAGATATTAAACAAGATTTCTGCTGCAATTCGG 780
Db |||||||
QY 3720 TATTTCTCAGACTCAAGTACCAATGTCAGAGATATTAAACAAGATTTCTGCTGCAATTCGG 3779
Db |||||||
QY 781 GCAATCCCTGTGTCATGTCATGTCCTTAGACACCTCAGTTCATTTGTTGCTTGGGCT 840
Db |||||||

3780 GCATGCGCCCTGTGCATGCGCATGTCCTAGACACCTCAGTTCAATGTGGTCTTGGCGCT 3839
841 TCTCTCTCTAGCAGCACCTCTCTGCTTCACTTAACCTCTGATGTCTTCACTCTCCTG 900
3840 TCTCTCTCTAGCAGCACCTCTCTGCTTCACTTAACCTCTGATGTCTTCACTCTCCTG 3899
901 CCAGCAACCCCAACCCCAAGTGCCTTCAAGGAGATAAATATCAATGGAACCTCAGAGATGAA 960
3900 CCAGCAACCCCAACCCCAAGTGCCTTCAAGGAGATAAATATCAATGGAACCTCAGAGATGAA 3959
961 CATCTAACCCACCTAGAGGAAACCCAGTTTGGTGATATAGAGCTTTATGTGGAGTGA AAA 1020
3960 CATCTAACCCACCTAGAGGAAACCCAGTTTGGTGATATAGAGCTTTATGTGGAGTGA AAA 4019
1021 TTGGGCAATGCCATTAATCTGCTTTCTTCTGTTTGTGTTTAAAGAAATGACGTTTACATATA 1080
4020 TTGGGCAATGCCATTAATCTGCTTTCTTCTGTTTGTGTTTAAAGAAATGACGTTTACATATA 4079
1081 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
4080 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4139
1141 CAATTAATGATAAATTAAGCATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
4140 CAATTAATGATAAATTAAGCATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4199
1201 CAATTAATGATAAATTAAGCATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
4200 CAATTAATGATAAATTAAGCATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4259
1261 RACAGGGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320
4260 RACAGGGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4319
1321 GCATAGCAAAATGACATAGCTACCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
4320 GCATAGCAAAATGACATAGCTACCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4379
1381 ATAAACGAATCTCTCTTGTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
4380 ATAAACGAATCTCTCTTGTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4439
1441 TCATCAAGAAACAAATCTCTTGTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
4440 TCATCAAGAAACAAATCTCTTGTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4499
1501 AAAACCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
4500 AAAACCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4559
1561 GCCGTGTGGCGTGGGAGGCTCCTGCTGAGCTTCCCTCCCAAGCTTTGCTGCTGAG 1620
4560 GCCGTGTGGCGTGGGAGGCTCCTGCTGAGCTTCCCTCCCAAGCTTTGCTGCTGAG 4619
1621 AGGACCAAGACAGCAGCAGCGCGGAAAGGCGCATCTAACCGTATCTAGGCTTTGG 1680
4620 AGGACCAAGACAGCAGCAGCGCGGAAAGGCGCATCTAACCGTATCTAGGCTTTGG 4679
1681 TAACTGCGGACAGTGTCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTTCCAGTTA 1740
4680 TAACTGCGGACAGTGTCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTTCCAGTTA 4739
1741 TAAATATTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
4740 TAAATATTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4799
1801 TATTTTAAATGCTGTAGTAAACATATGATTAATTTCTAGAAACAAACATCTAATA 1860
4800 TATTTTAAATGCTGTAGTAAACATATGATTAATTTCTAGAAACAAACATCTAATA 4859
1861 AGTATATAATCCTGTGAATAATGAGGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
4860 AGTATATAATCCTGTGAATAATGAGGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT

1921 CTAGAACCTGTAAACAGATAATAGAGAGTATTAATGATGGAACCTTAATAT 1980
4920 CTAGAACCTGTAAACAGATAATAGAGAGTATTAATGATGGAACCTTAATAT 4979
1981 ATAAATGCTGCCAGCGATTTTGTCTCAATATTTGTACTCTATCTATCTCTGCTGATATGG 2040
4980 ATAAATGCTGCCAGCGATTTTGTCTCAATATTTGTACTCTATCTATCTCTGCTGATATGG 5039
2041 AATTTCTTTTAAATTCAAACGCTGAAACGAAATCAGCATTTAGTCTTCCAGGACACCCCAA 2100
5040 AATTTCTTTTAAATTCAAACGCTGAAACGAAATCAGCATTTAGTCTTCCAGGACACCCCAA 5099
2101 TAATCAGTCAATGCTAAATGACAAAGTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTG 2160
5100 TAATCAGTCAATGCTAAATGACAAAGTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTG 5159
2161 TTTTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTG 5220
5160 TTTTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTG 5279
2221 TAAGGGTCTTGTAAAGAGAGTCTCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2280
5220 TAAGGGTCTTGTAAAGAGAGTCTCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 5279
2281 CCCTTTCTTGTAAATGATCATCATCTTCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 2340
5280 CCCTTTCTTGTAAATGATCATCATCTTCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 5339
2341 CTTAATGTACACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
5340 CTTAATGTACACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5399
2401 TTGTGTTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 2460
5400 TTGTGTTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 5459
2461 GAAGGTTGGGATCTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
5460 GAAGGTTGGGATCTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5519
2521 CTGTGTTGCTGGGTCACCGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
5520 CTGTGTTGCTGGGTCACCGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5579
2581 TGTACCATTAATTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
5580 TGTACCATTAATTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5639
2641 TGAAACGTCATGCTTTTGTGCAAAAGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700
5640 TGAAACGTCATGCTTTTGTGCAAAAGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5699
2701 AATTTCCACCTGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
5700 AATTTCCACCTGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2785
2761 AATTAATTTGTAAGGTTTCT 2785
5760 AATTAATTTGTAAGGTTTCT 5784

RESULT 9

ADD31094

ID ADD31094 standard; cDNA; 5784 BP.

XX

AC ADD31094;

XX

DT 15-JAN-2004 (first entry)

XX

DE Human cDNA encoding thrombospondin 2, THBS2.

XX

KW Human; thrombospondin 2; THBS2; ss; gene; single nucleotide polymorphism;

SNP: antiarteriosclerotic; cardiant; vasotropic; cerebroprotective;
cardiovascular-Gen; Thrombolytic; gene therapy; atherosclerosis;
coronary heart disease; myocardial infarction; stroke;
peripheral vascular diseases; venous thromboembolism; pulmonary embolism;
chromosome 6q27.

Homo sapiens.

Key Location/Qualifiers
CDS 240..3758
/*tag= b
/product= "Thrombospondin 2"
sig_peptide 240..293
/*tag= a
mat_peptide 294..3754
/*tag= c
/*tag= c
/note= "Mature thrombospondin 2"
replace(3949,G)
/*tag= d
/standard_name= "Single nucleotide polymorphism"
variation replace(5489,T)
/*tag= e
/standard_name= "Single nucleotide polymorphism"

US2003194703-A1.

16-OCT-2003.

13-NOV-2001; 2001US-00007781.

13-NOV-2000; 2000US-0248130P.

22-JUN-2001; 2001US-0300158P.

(BOLK/) BOLK S.

(DALEY/) DALEY G Q.

(MCCA/) MCCARTHY J J.

Bolk S, Daley GO, McCarthy JJ;

WPI; 2003-844447/78.

P-PSDB; ADD31095.

Predicting the likelihood of a vascular disease (e.g. stroke) in an individual comprises obtaining a nucleic acid sample from the individual and detecting a single nucleotide polymorphisms in the thrombospondin-2 gene.

Claim 2; SEQ ID NO 1; 17pp; English.

The invention relates to predicting the likelihood of a vascular disease, comprising obtaining a nucleic acid sample from the individual, and determining the genotype of the individual at nucleotide position 3949 of the thrombospondin-2 (THBS2) gene, where an individual who is homozygous for the variant allele has a decreased likelihood of a vascular disease as compared with an individual who is heterozygous or homozygous for the reference allele. Also included are diagnosing or aiding in the diagnosis of a vascular disease in an individual (comprising obtaining a nucleic acid sample from the individual, and determining the nucleotide present at nucleotide position 3949 of the thrombospondin-2 gene, where the presence of a T at nucleotide 3949 is indicative of an increased likelihood of a vascular disease in the individual, as compared with an individual having G at position 3949), a nucleic acid molecule comprising all or a portion of a sequence appearing as ADD31094 (where the nucleic acid molecule is at least 10 nucleotides in length and where the nucleic acid sequence comprises a polymorphic site at nucleotide position 3949 of ADD31094) and an allele-specific oligonucleotide that hybridizes to the nucleic acid molecule cited above. The composition and methods are useful in diagnosing, treating or predicting the clinical course and treatment response of vascular diseases (e.g. atherosclerosis, coronary heart disease, myocardial infarction, stroke, peripheral vascular diseases, venous thromboembolism or pulmonary embolism). The nucleic acid may also be used in developing new treatments for vascular disease and in developing cell culture-based and animal models for research and

CC treatment of the disease. The gene for THBS2 is located on chromosome
CC 6q27. The present sequence is the cDNA encoding thrombospondin 2.
XX
SQ Sequence 5784 bp; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;
Query Match 100.0%; Score 2785; DB 10; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTGGACGGTGTGATGACCGGGGTGATATTGTAAAGATGATTTTGAATGACAAC 60
DB |||||
QY 3000 GAGGACTTGGACGGTGTGATGACCGGGGTGATATTGTAAAGATGATTTTGAATGACAAC 3059
DB |||||
QY 61 ATCCAGATATTGATGATGATGTCCTGTAAGAAACAATGCCATCAGTGAGACAGATTGAGG 120
DB |||||
QY 3060 ATCCAGATATTGATGATGATGTCCTGTAAGAAACAATGCCATCAGTGAGACAGATTGAGG 3119
DB |||||
QY 121 AACTTCAGATGTCCTTGGATCCCAAGGACACCAAAATTTGATCCCACTGGTGC 180
DB |||||
QY 3120 AACTTCAGATGTCCTTGGATCCCAAGGACACCAAAATTTGATCCCACTGGTGC 3179
DB |||||
QY 181 ATTCCGCATCAAGCAAGGAGCTGGTTTCAGACAGCCAACTCGGACCCCGCATCGCTGTA 240
DB |||||
QY 3180 ATTCCGCATCAAGCAAGGAGCTGGTTTCAGACAGCCAACTCGGACCCCGCATCGCTGTA 3239
DB |||||
QY 241 GGTTCGACGATGTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 300
DB |||||
QY 3240 GGTTCGACGATGTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3299
DB |||||
QY 301 GACGACGATGTCCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTT 360
DB |||||
QY 3300 GACGACGATGTCCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTT 3359
DB |||||
QY 361 ATGTGGAAGCAGGTGACGACGACCTACTTGGGAGGACCAAGCCACCGCGGCTTATGGCTAC 420
DB |||||
QY 3360 ATGTGGAAGCAGGTGACGACGACCTACTTGGGAGGACCAAGCCACCGCGGCTTATGGCTAC 3419
DB |||||
QY 421 TCCGGCGTGTCCCTCAAGGTGGTGAATCTCCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB |||||
QY 3420 TCCGGCGTGTCCCTCAAGGTGGTGAATCTCCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3479
DB |||||
QY 481 GCGTGTGGACACGCGGGAACACGCGGGGAGGTGCGAACCTTATGCAACGACCCCAAGG 540
DB |||||
QY 3480 GCGTGTGGACACGCGGGAACACGCGGGGAGGTGCGAACCTTATGCAACGACCCCAAGG 3539
DB |||||
QY 541 AACATTGCTGGAAGGACTACACGCGCTATAGTGGACCTGACTCAGAGGCCCAAGACC 600
DB |||||
QY 3540 AACATTGCTGGAAGGACTACACGCGCTATAGTGGACCTGACTCAGAGGCCCAAGACC 3599
DB |||||
QY 601 GGCTACATCAGAGTCTTAGTGCATGAGGAAACAGGTTCATGGCAGACTCAGGACTATC 660
DB |||||
QY 3600 GGCTACATCAGAGTCTTAGTGCATGAGGAAACAGGTTCATGGCAGACTCAGGACTATC 3659
DB |||||
QY 661 TATGACCAAACTACGCTGGCGGGGCTGGGTCTATTGTCTTCTCTCAAGAAATGGTCT 720
DB |||||
QY 3660 TATGACCAAACTACGCTGGCGGGGCTGGGTCTATTGTCTTCTCTCAAGAAATGGTCT 3719
DB |||||
QY 721 TATTCTTCAGACCTCAAGTACGAATTCAGAGATATTAAACAAGATTGCTGCATTTCCG 780
DB |||||
QY 3720 TATTCTTCAGACCTCAAGTACGAATTCAGAGATATTAAACAAGATTGCTGCATTTCCG 3779
DB |||||
QY 781 GCAATGCCCTGTGCGATGCGATGGTCCCTAGACACCTCAGTTCATTTGCTGCTCCGGCT 840
DB |||||
QY 3780 GCAATGCCCTGTGCGATGCGATGGTCCCTAGACACCTCAGTTCATTTGCTGCTCCGGCT 3839
DB |||||
QY 841 TCTCTCTCTAGCAGACCTCTCTGTCCTTGAACCTTAACCTTCTGATGGTCTTCCACCTCTCG 900
DB |||||
QY 3840 TCTCTCTCTAGCAGACCTCTCTGTCCTTGAACCTTAACCTTCTGATGGTCTTCCACCTCTCG 3899
DB |||||
QY 901 CCAGCAACCCCAAAACCCAAAGTGCCTTCAGAGATATAATCAATGGAATCAGAGATGAA 960
DB |||||
QY 3900 CCAGCAACCCCAAAACCCAAAGTGCCTTCAGAGATATAATCAATGGAATCAGAGATGAA 3959
DB |||||

Qy	961	CATCTAACCCCACTAGAGAAACCAAGTCTTTGGTGATATATGAGACTTTTATGTGGAGTGAAAA	1021
Db	3960	CATCTAACCCCACTAGAGAAACCAAGTCTTTGGTGATATATGAGACTTTTATGTGGAGTGAAAA	4019
Qy	1021	TTGGGCATGCCATTACATCTGCTTTTCTTGTTTGGTTTAAAAGAATGACGTTTACATATA	1080
Db	4020	TTGGGCATGCCATTACATCTGCTTTTCTTGTTTGGTTTAAAAGAATGACGTTTACATATA	4079
Qy	1081	AAATGTAAATTACTTATTGTATTTATGTGTATATGGAGTTGAAGGGAATACCTGTGCATAAG	1140
Db	4080	AAATGTAAATTACTTATTGTATTTATGTGTATATGGAGTTGAAGGGAATACCTGTGCATAAG	4139
Qy	1141	CCATTAATGATAAATTAAGCATGAAAAATATTCGTGAACCTACTTTTGGTGCTTTAAAGTTGT	1200
Db	4140	CCATTAATGATAAATTAAGCATGAAAAATATTCGTGAACCTACTTTTGGTGCTTTAAAGTTGT	4199
Qy	1201	CACATATCTTCGAATTAGAGTTTGCTCTACATGACACACAAATCCCGCTAAATTAATTAATA	1260
Db	4200	CACATATCTTCGAATTAGAGTTTGCTCTACATGACACACAAATCCCGCTAAATTAATTAATA	4259
Qy	1261	AACAAGGCTCAATTCAAATTTGAAAGTAATGTTTTAGTAAAGAGAGATTAGAAGACAACAG	1320
Db	4260	AACAAGGCTCAATTCAAATTTGAAAGTAATGTTTTAGTAAAGAGAGATTAGAAGACAACAG	4319
Qy	1321	GCATAGCAATGACATTAAGCTACCGATTAACTAATCGGAACATGTAAAAAGTTACAAAA	1380
Db	4320	GCATAGCAATGACATTAAGCTACCGATTAACTAATCGGAACATGTAAAAAGTTACAAAA	4379
Qy	1381	ATAAACGAACCTCCTCTGTGCTTACAAATGAAGCCCTCATGTGCAATGAGATGACAGTT	1440
Db	4380	ATAAACGAACCTCCTCTGTGCTTACAAATGAAGCCCTCATGTGCAATGAGATGACAGTT	4439
Qy	1441	TCATCAAGAAACCAACATCCTTGCMAATGGGTGTGACGCGGTTCCAGATGTGATTTGCG	1500
Db	4440	TCATCAAGAAACCAACATCCTTGCMAATGGGTGTGACGCGGTTCCAGATGTGATTTGCG	4499
Qy	1501	AAAACCTCATTTAAGTAAAAAGTTAGCAGAGCAAAAGTCGGTCTTTAGCTGCTGTTCT	1560
Db	4500	AAAACCTCATTTAAGTAAAAAGTTAGCAGAGCAAAAGTCGGTCTTTAGCTGCTGTTCT	4559
Qy	1561	GCGTGTGCGCTCGGGGGCTCCTGCCTGAGCTTCCCTCCAGGTTTGTCTGCTGCTGAG	1620
Db	4560	GCGTGTGCGCTCGGGGGCTCCTGCCTGAGCTTCCCTCCAGGTTTGTCTGCTGCTGAG	4619
Qy	1621	AGGAACACAGAGCAGACGCA CAGGCGGAAAGCGCATCTTAACGCGTATCTAGGCTTCGG	1680
Db	4620	AGGAACACAGAGCAGACGCA CAGGCGGAAAGCGCATCTTAACGCGTATCTAGGCTTCGG	4679
Qy	1681	TAACTCGGGCAAGTGTGCTTTTACCTGATTTGATGATACATTTCAATTAAGTTTCCAGTTA	1740
Db	4680	TAACTCGGGCAAGTGTGCTTTTACCTGATTTGATGATACATTTCAATTAAGTTTCCAGTTA	4739
Qy	1741	TAAATATTTTGTAAATTTATTAAGTGACTATAGAAATGCACTCCATTTACAGTAACT	1800
Db	4740	TAAATATTTTGTAAATTTATTAAGTGACTATAGAAATGCACTCCATTTACAGTAACT	4799
Qy	1801	TATTTTAAATPATGCCTAGTAAACATATGTAGTATAAATTTCTAGAAACAAACATCTAATA	1860
Db	4800	TATTTTAAATPATGCCTAGTAAACATATGTAGTATAAATTTCTAGAAACAAACATCTAATA	4859
Qy	1861	AGTATATATCCTGTGAAAATATGAGGCTTGATAATTTAGGTTGTACGATGAAGCATG	1920
Db	4860	AGTATATATCCTGTGAAAATATGAGGCTTGATAATTTAGGTTGTACGATGAAGCATG	4919
Qy	1921	CTAGAAGCTGTAACAGAAATATGAGGCTTGATAATTTAGGTTGTACGATGAAGCATG	4979
Db	4920	CTAGAAGCTGTAACAGAAATATGAGGCTTGATAATTTAGGTTGTACGATGAAGCATG	5039
Qy	1981	ATAATGTGTCCAGCGATTTTAGTTTCAATATTTGTTACTGTATCTATCTCTGCTATATGG	2040
Db	4980	ATAATGTGTCCAGCGATTTTAGTTTCAATATTTGTTACTGTATCTATCTCTGCTATATGG	5099
Qy	2041	AAATCTTTTAAATTCAAACGCTGAAACGAATCAGCATTTTAGTCTTTGCGACGACACCCAA	2100

[illegible]

RESULT 10	
ADE85079	
ID	ADE85079 standard; DNA; 5784 BP.
XX	
XX	
AC	ADE85079;
XX	
DT	29-JAN-2004 (first entry)
XX	
DE	Farnesyl transferase inhibitor modulated leukemia associated gene #298.
XX	
XX	ss; cytostatic; farnesyl transferase inhibitor; gene expression;
KW	quinolinone; leukemia; cancer.
KW	


```
QY 1681 TAACTGCGGACAGTGTCTTTTACCTGATTGATGATACATTTTCATTAAGGTTCCAGTTA 1740
Db |||||||
QY 4680 TAACTGCGGACAGTGTCTTTTACCTGATTGATGATACATTTTCATTAAGGTTCCAGTTA 4739
Db |||||||
QY 1741 TAAATATTTTGTATATTTATTAAGTGACTAGAACTCAATTTACAGTAACT 1800
Db |||||||
QY 4740 TAAATATTTTGTATATTTATTAAGTGACTAGAACTCAATTTACAGTAACT 4799
Db |||||||
QY 1801 TATTTTAAATATGCTTAGTAAACACATATGTAGTATATAATTTCTAGAAACAAACATCTTAATA 1860
Db |||||||
QY 4800 TATTTTAAATATGCTTAGTAAACACATATGTAGTATATAATTTCTAGAAACAAACATCTTAATA 4859
Db |||||||
QY 1861 AGTATATATCTCTGTGAATAATAGAGCTGTGATAATATTAGTGTGTGACGATGAAGCATG 1920
Db |||||||
QY 4860 AGTATATATCTCTGTGAATAATAGAGCTGTGATAATATTAGTGTGTGACGATGAAGCATG 4919
Db |||||||
QY 1921 CTAGAAGCTGTAAACAGATACATAGAGATATAGGAGTTTATGATGGAACCTTAATAT 1980
Db |||||||
QY 4920 CTAGAAGCTGTAAACAGATACATAGAGATATAGGAGTTTATGATGGAACCTTAATAT 4979
Db |||||||
QY 1981 ATAAATGTTGCCAGCGATTTTGTCTCAATATTTGTACTTATCTATCTCTGTATATATGG 2040
Db |||||||
QY 4980 ATAAATGTTGCCAGCGATTTTGTCTCAATATTTGTACTTATCTATCTCTGTATATATGG 5039
Db |||||||
QY 2041 AATTCCTTTTAAATCAACGCTGAACCGAATCAGCATTTAGTCTTCCAGGCACACCCAA 2100
Db |||||||
QY 5040 AATTCCTTTTAAATCAACGCTGAACCGAATCAGCATTTAGTCTTCCAGGCACACCCAA 5099
Db |||||||
QY 2101 TAATCAGTCATGTGTAATATGACAAAGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 2160
Db |||||||
QY 5100 TAATCAGTCATGTGTAATATGACAAAGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 5159
Db |||||||
QY 2161 TTTTTTGTCTTTAAGTTCGATGATCTTCTCTGAGGAATAGTCACTCATCCACCTCCACA 2220
Db |||||||
QY 5160 TTTTTTGTCTTTAAGTTCGATGATCTTCTCTGAGGAATAGTCACTCATCCACCTCCACA 5219
Db |||||||
QY 2221 TAAGGGTTTAGTAAGAGAAGTCTCTGCTGATGATGGATAGGGGCAAACTCTTTTC 2280
Db |||||||
QY 5220 TAAGGGTTTAGTAAGAGAAGTCTCTGCTGATGATGGATAGGGGCAAACTCTTTTC 5279
Db |||||||
QY 2281 CCCTTCTGTTTAAATAGTCATCACATTTCTATGCCAAACAGGAACGATCCATACTTTAGT 2340
Db |||||||
QY 5280 CCCTTCTGTTTAAATAGTCATCACATTTCTATGCCAAACAGGAACGATCCATACTTTAGT 5339
Db |||||||
QY 2341 CTTAATGTACACATTCATTTGTATAAATAATTTGTTGTTCTTCTGAGTGTGATCG 2400
Db |||||||
QY 5340 CTTAATGTACACATTCATTTGTATAAATAATTTGTTGTTCTTCTGAGTGTGATCG 5399
Db |||||||
QY 2401 TTGTGTTGTTTGTGCACTTTTACTTTTTCGTTGTTGAGCTGTATTCGCGAGACAAC 2460
Db |||||||
QY 5400 TTGTGTTGTTTGTGCACTTTTACTTTTTCGTTGTTGAGCTGTATTCGCGAGACAAC 5459
Db |||||||
QY 2461 GAACGTTGGGATACCTCATTAATATGACCACTGTCAACGCTGCAAGTTTCTGTTT 2520
Db |||||||
QY 5460 GAACGTTGGGATACCTCATTAATATGACCACTGTCAACGCTGCAAGTTTCTGTTT 5519
Db |||||||
QY 2521 CTGTGTTGTTGGGTCACACCGTACAAATGTTGGGAATGACCATGATGTGATATTTAGAA 2580
Db |||||||
QY 5520 CTGTGTTGTTGGGTCACACCGTACAAATGTTGGGAATGACCATGATGTGATATTTAGAA 5579
Db |||||||
QY 2581 TGTAACATATTTTGTGTAATATTTATGTTTTTCTAAACAAATTTATCGTATAGGTTGA 2640
Db |||||||
QY 5580 TGTAACATATTTTGTGTAATATTTATGTTTTTCTAAACAAATTTATCGTATAGGTTGA 5639
Db |||||||
QY 2641 TGAACGTCATGCTGTTTGGCAAGACGTGAATATTTATTTATGTTTCATCGTCAAGTCAA 2700
Db |||||||
QY 5640 TGAACGTCATGCTGTTTGGCAAGACGTGAATATTTATTTATGTTTCATCGTCAAGTCAA 5699
Db |||||||
QY 2701 AATTTTCAACCTGAAACCTGCACTTACGTAGAACCTCATTTTAAAGATTTAAACAGG 2760
Db |||||||
QY 5700 AATTTTCAACCTGAAACCTGCACTTACGTAGAACCTCATTTTAAAGATTTAAACAGG 5759
Db |||||||
QY 2761 AAATAAATTTGAAAAAGGTTTTCT 2785
```

Db 5760 AAATAAATTTGAAAAAGGTTTTCT 5784

RESULT 11

ADQ18777

ID ADQ18777 standard; DNA; 5784 BP.

XX ADQ18777;

XX 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 1596.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human; ds.

XX Homo sapiens.

XX WO2004048938-A2.

XX 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.

XX 26-NOV-2002; 2002US-0429739P.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Ginsburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

XX Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

XX Example 2; SEQ ID NO 1596; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated DNA of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

XX Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;

Query Match 100.0%; Score 2785; DB 12; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGACTTGCACCGTGTATGACGGGGTGATATTGTTAAAGATGATTTTGCAATGACAAC 60
Db |||||||
QY 3000 GAGGACTTGCACCGTGTATGACGGGGTGATATTGTTAAAGATGATTTTGCAATGACAAC 3059
Db |||||||
QY 61 ATCCAGATATTGATGATGTGTCTCTGAAACAATGCCATCAGTGAGACAGCTTCAGG 120
Db |||||||
QY 3060 ATCCAGATATTGATGATGTGTCTCTGAAACAATGCCATCAGTGAGACAGCTTCAGG 3119
Db |||||||
QY 121 AACTTCAGATGTCCCTTTGGATCCCAAGGGACCAACCCAAATTTGATCCCAACTGGGTC 180
Db |||||||
QY 3120 AACTTCAGATGTCCCTTTGGATCCCAAGGGACCAACCCAAATTTGATCCCAACTGGGTC 3179
Db |||||||
QY 181 ATTGCCATCAAGCAAGGAGCTGGTTTCAGACAGCCAACCTCGGACCCCGCATCGTGTGA 240

QY 1 GAGGACTTGGACGGGTGATGATATTTGTAAGATGATTTTGCAATGACAAC 60
DB |||||
2892 GAGGACTTGGACGGGTGATGACGGGTGATTTTGTAAAGATGATTTTGCAATGACAAC 2951
QY 61 ATCCAGATATTGATGATGTCGTCTGTAAGAAACAAATGCCATCAGTGAGACAGACTTCAGG 120
DB |||||
2952 ATCCAGATATTGATGATGTCGTCTGTAAGAAACAAATGCCATCAGTGAGACAGACTTCAGG 3011
QY 121 AACTTCCAGATGTCCTCCCTTGGATCCCAAGGACCAACCCAAATTCGATCCCACTGGGTC 180
DB |||||
3012 AACTTCCAGATGTCCTCCCTTGGATCCCAAGGACCAACCCAAATTCGATCCCACTGGGTC 3071
QY 181 ATTGCGCATCAAGGCAAGAGCTGGTTTCAGACAGCCAACTCGGACCCCGCATCGCTGTA 240
DB |||||
3072 ATTGCGCATCAAGGCAAGAGCTGGTTTCAGACAGCCAACTCGGACCCCGCATCGCTGTA 3131
QY 241 GGTTTTGGACGAGTTTGGGTCGTGGACTTCAGTGGGACACATTTCTACGTAACACTGACCCG 300
DB |||||
3132 GGTTTTGGACGAGTTTGGGTCGTGGACTTCAGTGGGACACATTTCTACGTAACACTGACCCG 3191
QY 301 GACGACGACTATGCTGGCTTCGTCTTTGGTTTACAGTCAAGCAGCCGCTTCTATGTTGGTG 360
DB |||||
3192 GACGACGACTATGCTGGCTTCGTCTTTGGTTTACAGTCAAGCAGCCGCTTCTATGTTGGTG 3251
QY 361 ATGTGGAAGCAGGTGACGACGACCTACTGGGAGGACCAAGCCCAACGCGGCTTATGGCTAC 420
DB |||||
3252 ATGTGGAAGCAGGTGACGACGACCTACTGGGAGGACCAAGCCCAACGCGGCTTATGGCTAC 3311
QY 421 TCCGGCTGTCTCCCTCAAGGTGTGAACTCCACACGGGACCGGCGAGCACCTGAGGAAC 480
DB |||||
3312 TCCGGCTGTCTCCCTCAAGGTGTGAACTCCACACGGGACCGGCGAGCACCTGAGGAAC 3371
QY 481 GCGCTGTGGCAACACGGGGAAACGCGCGGGCAGGTGCGAACTTATGGCACGACCCACGG 540
DB |||||
3372 GCGCTGTGGCAACACGGGGAAACGCGCGGGCAGGTGCGAACTTATGGCACGACCCACGG 3431
QY 541 AACATTGGCTGGAAGGACTACACGGCTATAGGTGGCACTGACTCAGAGCCCAAGACC 600
DB |||||
3432 AACATTGGCTGGAAGGACTACACGGCTATAGGTGGCACTGACTCAGAGCCCAAGACT 3491
QY 601 GGCTACATCAGAGTCTTAGTGATCAAGGAAACACAGGTATGGCAGACTCAGGACCTATC 660
DB |||||
3492 GGCTACATCAGAGTCTTAGTGATCAAGGAAACACAGGTATGGCAGACTCAGGACCTATC 3551
QY 661 TATGACCAAACTCAGCTGGCGGGCGGTGGTCTATTGTCCTCTCAAGAAATGTC 720
DB |||||
3552 TATGACCAAACTCAGCTGGCGGGCGGTGGTCTATTGTCCTCTCAAGAAATGTC 3611
QY 721 TATTTCTCAGACCTCAAGTACGAATGCGAGAGATATTTAAACAAGATTTGCTGCAATTCGG 780
DB |||||
3612 TATTTCTCAGACCTCAAGTACGAATGCGAGAGATATTTAAACAAGATTTGCTGCAATTCGG 3671
QY 781 GCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTTCATTTGCTTGGCTTCGGCT 840
DB |||||
3672 GCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTTCATTTGCTTGGCTTCGGCT 3731
QY 841 TCTCTCTCAGACCACTCCTGTCCTCTGACCTTAACTCTGATGGTTCTTCACTCCTG 900
DB |||||
3732 TCTCTCTCAGACCACTCCTGTCCTCTGACCTTAACTCTGATGGTTCTTCACTCCTG 3791
QY 901 CCAGAACCCCAACCCAAAGTGGCTTCAGAGATAAATATCAATGGAACCTCAGAGATGAA 960
DB |||||
3792 CCAGAACCCCAACCCAAAGTGGCTTCAGAGATAAATATCAATGGAACCTCAGAGATGAA 3851
QY 961 CATCTAACCCCACTAGAGAAACCAAGTTTGGTGATATAGACACTTATGTGGAGTGAATA 1020
DB |||||
3852 CATCTAACCCCACTAGAGAAACCAAGTTTGGTGATATAGACACTTATGTGGAGTGAATA 3911
QY 1021 TTGGGCACTGCAATACATTTGCTTTTCTGTTGTTTAAAGAAATGACGTTTACATATA 1080
DB |||||
3912 TTGGGCACTGCAATACATTTGCTTTCTGTTGTTTAAAGAAATGACGTTTACATATA 3971
QY 1081 AAATGTAATTAATTAATGTTATGTTATGAGAGTTGAAGGGAATPACTGTGCATAAG 1140

DB |||||
3972 AAATGTAATTAATTAATGTTATGTTATGTTATGAGTTGAAGGGAATACCTGTGCATAAG 4031
QY 1141 CCATTATGATAAATTAAGCATGAAAAATATGCTGAACACTACTTTTGGTGTCTAAAGTTGT 1200
DB |||||
4032 CCATTATGATAAATTAAGCATGAAAAATATGCTGAACACTACTTTTGGTGTCTAAAGTTGT 4091
QY 1201 CACTATTCTTGAATTAAGTGTCTTACAAATGACACAAAAATCCCGCTAAATTAATATA 1260
DB |||||
4092 CACTATTCTTGAATTAAGTGTCTTACAAATGACACAAAAATCCCGCTAAATTAATATA 4151
QY 1261 AACCAAGGCTCAATTCAAATTTGAAGTAATGTTTATAGTAGGAGAGATTAGACACACAG 1320
DB |||||
4152 AACCAAGGCTCAATTCAAATTTGAAGTAATGTTTATAGTAGGAGAGATTAGACACACAG 4211
QY 1321 GCATAGCAATGACATAAGCTACCGATTAACTAACTCGGAACATGTAAAAACAGTTACAAA 1380
DB |||||
4212 GCATAGCAATGACATAAGCTACCGATTAACTAACTCGGAACATGTAAAAACAGTTACAAA 4271
QY 1381 ATAAAGCACTCTCTCTTGTCTTACATGAAAGCCCTCATGTGCAGTAGAGATGAGTT 1440
DB |||||
4272 ATAAAGCACTCTCTCTTGTCTTACATGAAAGCCCTCATGTGCAGTAGAGATGAGTT 4331
QY 1441 TCATCAAAAGAACAAACATCCTTGTCTTACATGAAAGCCCTCATGTGCAGTAGAGATGAGTT 1500
DB |||||
4332 TCATCAAAAGAACAAACATCCTTGTCTTACATGAAAGCCCTCATGTGCAGTAGAGATGAGTT 4391
QY 1501 AAAACCTCATTTAAGTAAAGGTTAGCAGAGCAAAAGTCGGTGTCTTGTAGCTGTCTGTCT 1560
DB |||||
4392 AAAACCTCATTTAAGTAAAGGTTAGCAGAGCAAAAGTCGGTGTCTTGTAGCTGTCTGTCT 4451
QY 1561 GCGGTTGTGGCTCGGAGGCTCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG 1620
DB |||||
4452 GCGGTTGTGGCTCGGAGGCTCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG 4511
QY 1621 AGGAACACAGAGCAGACGACAGCGGAAAGGCGCATCTAACCGTATCTAGGCTTTGG 1680
DB |||||
4512 AGGAACACAGAGCAGACGACAGCGGAAAGGCGCATCTAACCGTATCTAGGCTTTGG 4571
QY 1681 TAACCTGCGGACAGAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 1740
DB |||||
4572 TAACCTGCGGACAGAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 4631
QY 1741 TAAATATTTTGTAAATATTTTAAAGTGAATAGAAATGCAATCCATTTTACAGTAAT 1800
DB |||||
4632 TAAATATTTTGTAAATATTTTAAAGTGAATAGAAATGCAATCCATTTTACAGTAAT 4691
QY 1801 TATTTAAATATGCTAGTAAACATATGAGGCTTGATTAATTTAGGTTGTACAGATGAAGCATG 1860
DB |||||
4692 TATTTAAATATGCTAGTAAACATATGAGGCTTGATTAATTTAGGTTGTACAGATGAAGCATG 4751
QY 1861 AGTATATATCTGTGAAATATGAGGCTTGATTAATTTAGGTTGTACAGATGAAGCATG 1920
DB |||||
4752 AGTATATATCTGTGAAATATGAGGCTTGATTAATTTAGGTTGTACAGATGAAGCATG 4811
QY 1921 CTAGAGCTGTAAACAGATACATAGAGATTAATGAGGAGTTTATGATGGAACCTT-AAATA 1979
DB |||||
4812 CTAGAGCTGTAAACAGATACATAGAGATTAATGAGGAGTTTATGATGGAACCTTAAATA 4871
QY 1980 TATAATGTTGCGACGAGTTTGTAGTTCAATTTTGTACTGTTATCTATCTGCTGTATATG 2039
DB |||||
4872 TATAATGTTGCGACGAGTTTGTAGTTCAATTTTGTACTGTTATCTATCTGCTGTATATG 4931
QY 2040 GAATCTTTTAAATCAACGCTGAAAAACGAATCAGCAATTTAGTCTTCCAGGACACACCA 2099
DB |||||
4932 GAATCTTTTAAATCAACGCTGAAAA- GAATCAGCAATTTAGTCTTCCAGGACACACCA 4990
QY 2100 ATATCAGTCACTGTGTAATGACAAAGTTTGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTG 2159
DB |||||
4991 ATATCAGTCACTGTGTAATGACAAAGTTTGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTG 5050
QY 2160 ---GTTTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTG 2215

QY 601 GGTACATCAGAGTCTTAGTGATGAAGGAAAAACAGGTCATGGCAGACTCAGGACCTATC 660
Db 3492 GGTACATCAGAGTCTTAGTGATGAAGGAAAAACAGGTCATGGCAGACTCAGGACCTATC 3551
QY 661 TATGACCAACCTACGCTGGCGGGCGCTGGGCTATTGTCTCTCAAGAAATGGTC 720
Db 3552 TATGACCAACCTACGCTGGCGGGCGCTGGGCTATTGTCTCTCAAGAAATGGTC 3611
QY 721 TATTTCTCAGACCTCAAGTAGCAAGTACAGAGATATTAAACAAGATTGTGCTGCAATTCGG 780
Db 3612 TATTTCTCAGACCTCAAGTAGCAAGTACAGAGATATTAAACAAGATTGTGCTGCAATTCGG 3671
QY 781 GCAGTGGCTGTGCATGCCATGTGCTCCTAGACACCTCAGTTCATTTGTGGTCTTGGCGGT 840
Db 3672 GCAGTGGCTGTGCATGCCATGTGCTCCTAGACACCTCAGTTCATTTGTGGTCTTGGCGGT 3731
QY 841 TCTCTCTCTAGCAGCACCTCTGCTGCTTGCCTTAACCTCAGTTCATTTGTGGTCTTGGCGGT 900
Db 3732 TCTCTCTCTAGCAGCACCTCTGCTGCTTGCCTTAACCTCAGTTCATTTGTGGTCTTGGCGGT 3791
QY 901 CCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960
Db 3792 CCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 3851
QY 961 CATCTAACCCACTAGAGGAAACAGTTTGGTGATATCAGACTTTTATGTGGAGTGAAAA 1020
Db 3852 CATCTAACCCACTAGAGGAAACAGTTTGGTGATATCAGACTTTTATGTGGAGTGAAAA 3911
QY 1021 TTGGGATGCCATTAATCATTTGCTTTCTGTTGTTTAAAGAAATGAGCTTTACATATA 1080
Db 3912 TTGGGATGCCATTAATCATTTGCTTTCTGTTGTTTAAAGAAATGAGCTTTACATATA 3971
QY 1081 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
Db 3972 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4031
QY 1141 CCATTAATGATAAATTAAGCATGAAAAATATGCTGAACTACTTTTGTGCTTAAAGTTGT 1200
Db 4032 CCATTAATGATAAATTAAGCATGAAAAATATGCTGAACTACTTTTGTGCTTAAAGTTGT 4091
QY 1201 CACTATTCTTGAATTAAGCATGAAAAATATGCTGAACTACTTTTGTGCTTAAAGTTATA 1260
Db 4092 CACTATTCTTGAATTAAGCATGAAAAATATGCTGAACTACTTTTGTGCTTAAAGTTATA 4151
QY 1261 AACAGGGTCAATTCAAATTTGAAGTAATGTTTGTAGTAAGGAGATTTAGAGCAACAG 1320
Db 4152 AACAGGGTCAATTCAAATTTGAAGTAATGTTTGTAGTAAGGAGATTTAGAGCAACAG 4211
QY 1321 GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAACAGTTACAAA 1380
Db 4212 GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAACAGTTACAAA 4271
QY 1381 ATAAACGAATCTCCTCTGTCTACATGAAAGCCCTCATGTGAGTAGAGATGCAAGTT 1440
Db 4272 ATAAACGAATCTCCTCTGTCTACATGAAAGCCCTCATGTGAGTAGAGATGCAAGTT 4331
QY 1441 TCATCAAGAACAAACATCTTGTCAAAATGGGTGTGACGCGTTCACAGATGTTGGC 1500
Db 4332 TCATCAAGAACAAACATCTTGTCAAAATGGGTGTGACGCGTTCACAGATGTTGGC 4391
QY 1501 AAAACCTCATTTAAGTAAAGGTTAGCAGACCAAGTGCCTGCTTTAGCTGCTGCTTGT 1560
Db 4392 AAAACCTCATTTAAGTAAAGGTTAGCAGACCAAGTGCCTGCTTTAGCTGCTGCTTGT 4451
QY 1561 GCCGTGTGGCGTGGGGAGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG 1620
Db 4452 GCCGTGTGGGTGGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG 4511
QY 1621 AGGAACAGAGCAGCAGCAGGCGGAAAGGCGCATCTAAAGCGTATCTAGGCTTTGG 1680
Db 4512 AGGAACAGAGCAGCAGCAGGCGGAAAGGCGCATCTAAAGCGTATCTAGGCTTTGG 4571

QY 1681 TAACTCGGCAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 1740
Db 4572 TAACTCGGCAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 4631
QY 1741 TAAATATTTTGTAAATATTTAATTAAGTACTATAGAAATGCAACTCCATTTTACCAAGTAAC 1800
Db 4632 TAAATATTTTGTAAATATTTAATTAAGTACTATAGAAATGCAACTCCATTTTACCAAGTAAC 4691
QY 1801 TATTTTAAATATGCTCTAGTAACACATATGTAGTATAAATTTCTAGAAACAAACATCTAATA 1860
Db 4692 TATTTTAAATATGCTCTAGTAACACATATGTAGTATAAATTTCTAGAAACAAACATCTAATA 4751
QY 1861 AGTATATATCTCTGABAAATATAGGCTTGATTAATTAATTAAGTGTGACATGAAGCATG 1920
Db 4752 AGTATATATCTCTGABAAATATAGGCTTGATTAATTAATTAAGTGTGACATGAAGCATG 4811
QY 1921 CTAGAAGCTGTAAACAGAAATACATAGAGAAATATAGGAGTATTAATGATGAACCTTAAATA 1979
Db 4812 CTAGAAGCTGTAAACAGAAATACATAGAGAAATATAGGAGTATTAATGATGAACCTTAAATA 4871
QY 1980 TATAATGTTGCCAGCGATTTTATGTTCAATATTTGTTACTTCTATCTCTGCTGTATATG 2039
Db 4872 TATAATGTTGCCAGCGATTTTATGTTCAATATTTGTTACTTCTATCTCTGCTGTATATG 4931
QY 2040 GAAATCTTTTAAATCAAAAGCTGAAAAAGAAATCAGCAATTTAGTCTTGCCAGGACACCCCA 2099
Db 4932 GAAATCTTTTAAATCAAAAGCTGAAAAAGAAATCAGCAATTTAGTCTTGCCAGGACACCCCA 4990
QY 2100 ATAAATCAGTCATGTGTAATATGCAAAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2159
Db 4991 ATAAATCAGTCATGTGTAATATGCAAAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5050
QY 2160 ---GTTTGTGTTTAAAGTTGCAATGATCTTCTGAGGAAATAGTCACTCATCCACT 2215
Db 5051 GTTGTGTTTGTGTTTAAAGTTGCAATGATCTTCTGAGGAAATAGTCACTCATCCACT 5110
QY 2216 CCACATAGGGGTTTATGTAAGAGAGTCTGCTGCTGATGATGATAGGGGCAAAATCT 2275
Db 5111 CCACATAGGGGTTTATGTAAGAGAGTCTGCTGCTGATGATGATAGGGGCAAAATCT 5170
QY 2276 TTTTCCCTCTTCTGTTAATAGTCATACATTTCTATGCCCACCAAGAGACGATCCATACT 2335
Db 5171 TTTTCCCTCTTCTGTTAATAGTCATACATTTCTATGCCCACCAAGAGACGATCCATACT 5230
QY 2336 TTAGCTTTAATGTCACATGCAATTTTATGATAAATTAATTTTGTGTTTCTCTTGAGGTT 2395
Db 5231 TTAGCTTTAATGTCACATGCAATTTTATGATAAATTAATTTTGTGTTTCTCTTGAGGTT 5290
QY 2396 GATCGTTG---TGTTGTTTGTGCTGCACTTTTACTTTTTCGGTGTGAGCTGATTTCCC 2452
Db 5291 GATCGTTGTTGTTGTTGCTGCACTTTTACTTTTCTTTTTCGGTGTGAGCTGATTTCCC 5350
QY 2453 GAGA-CAACGAAGCGTTGGGATCTCTCAATTAATGTTAGCGACTCTCAACAGCGTGCAGGT 2511
Db 5351 GAGACCAACGAGCGTTGGGATCTCTCAATTAATGTTAGCGACTCTCAACAGCGTGCAGGT 5410
QY 2512 TTTCTGTTCTGTTGTGGGTCACCGTCAATGTTGGGAAATGACGATGATGAA 2571
Db 5411 TTTCTGTTCTGTTGTGGGTCACCGTCAATGTTGGGAAATGACGATGATGAA 5470
QY 2572 TATTTAGAAATGACCAATTTTGTGTAATTTATGTTTCTTAAACAAATTTATCGT 2631
Db 5471 TATTTAGAAATGACCAATTTTGTGTAATTTATGTTTCTTAAACAAATTTATCGT 5530
QY 2632 ATAGGTTGATGAAACGTCATGTTGTTTCCAAAGACTGTAAATATTTATGTTGTTCA 2691
Db 5531 ATAGGTTGATGAAACGTCATGTTGTTTCCAAAGACTGTAAATATTTATGTTGTTCA 5590
QY 2692 CATGGTCAAAATTTCCACTGAACCCCTGCACTTAGCTAGAACCTCATTTTAAAGATT 2751
Db 5591 CATGGTCAAAATTTCCACTGAACCCCTGCACTTAGCTAGAACCTCATTTTAAAGATT 5650
QY 2752 AACCAAGGAAATAAATTTGTAAGGTTTCTTCT 2785

us-10-017-724-1_copy_3000_5784.rng

Db 5651 AACACACGGGAAATAATTGTAATAAGGTTTCT 5684

Search completed: November 22, 2004, 17:22:57
Job time : 1248.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:21 ; Search time 11525.2 Seconds
(without alignments)
11427.340 Million cell updates/sec

Title: US-10-017-724-1_COPY_3000_5784

Perfect score: 2785
Sequence: 1 gaggactggacgtgatgg.....aattgtaaaaagggtttct 2785

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2785	100.0	5784	6	AX330488 Sequence
2	2785	100.0	5784	6	AX375587 Sequence
3	2785	100.0	5784	6	AX409513 Sequence
4	2785	100.0	5784	6	AX480837 Sequence
5	2785	100.0	5784	6	AX774982 Sequence
6	2785	100.0	5784	9	HUMTHRSPO
7	2703	97.1	3339	9	HS807201
8	2372.8	85.2	5505	11	BV177227
9	2372.8	85.2	5505	11	BV177983
10	2372.8	85.2	5505	11	BV178635
11	2372.8	85.2	5505	11	BV179851
12	1985	71.3	2026	11	G06722
13	1963.8	70.5	159175	9	BX322234
14	1694	60.8	2780	9	HUMTHRSPD
15	809	29.0	921	11	BV177085
16	809	29.0	921	11	BV177192
17	809	29.0	921	11	BV177656
18	793.8	28.5	2154	10	BC031843
19	783.8	28.1	4125	10	BC053702

20	767.2	27.5	4697	4	BTTHROM2
21	753.6	27.1	4108	6	BD129562
22	753.6	27.1	4108	10	MUSTSF2B
23	739.8	26.6	2764	6	CQ714089
24	687.2	24.7	814	11	BV166806
25	595.6	21.4	619	6	BD265321
26	595.6	21.4	619	6	AR401307
27	595.6	21.4	619	6	AX192760
28	594.8	21.4	848	6	BD216525
29	585.8	21.0	928	4	AF454919
30	540.2	19.4	637	6	AX261797
31	531.8	19.1	3537	5	CHKTHBSA
32	516	18.5	516	6	AX397018
33	497.6	17.9	561	11	BV167541
34	495.4	17.8	5269	4	AB005287
35	495.4	17.8	5289	6	BD129598
36	492.2	17.7	2162	4	BTRNAT1
37	491.8	17.7	7231	6	AR447356
38	491.8	17.7	7232	6	AX281758
39	490.2	17.6	4434	6	BD129560
40	490.2	17.6	4434	9	HSTHROMR
41	490.2	17.6	5719	6	CQ727671
42	490.2	17.6	5722	6	CQ776339
43	490.2	17.6	5722	6	AX094823
44	490.2	17.6	5722	6	AX411057
45	490.2	17.6	5722	6	AX587747

ALIGNMENTS

RESULT 1
AX330488
LOCUS AX330488
DEFINITION Sequence 997 from Patent WO0194629.
ACCESSION AX330488
VERSION AX330488.1 GI:18103466
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 997 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
source
1..5784
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 2785; DB 6; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAGGACTTGGACGCGTGATGGACGGGGTGATATTGTAAGATGATTTTGACAAATGACAAAC	60
DB	3000	GAGGACTTGGACGCGTGATGGACGGGGTGATATTGTAAGATGATTTTGACAAATGACAAAC	3059
QY	61	ATCCCGATATTGATGATGTGTCTCTGAAACAAATGCCATCAGTGACAGACTTCAGG	120
DB	3060	ATCCCGATATTGATGATGTGTCTCTGAAACAAATGCCATCAGTGACAGACTTCAGG	3119
QY	121	AACCTCCAGATGGTCCCTTGGATCCCAAGGGACCAACCAAAATTTGATCCCAACTGGGTC	180
DB	3120	AACCTCCAGATGGTCCCTTGGATCCCAAGGGACCAACCAAAATTTGATCCCAACTGGGTC	3179
QY	181	ATTCGCCATCAGGACGAGCTGGTTTCAGACACCACTCGGACCCCGGATCGCTGTA	240

Db 3180 ATTGCGCATCAGGCAAGAGCTGGTTCACAGACCACTCGACCCCGCATCGCTGTA 3239
Qy 241 GGTTTTCAAGAGTTGGTCTGTGGCACTTCAGTGGCACTTCAGTAAACACATGACCGG 300
Db 3240 GGTTTTCAAGAGTTGGTCTGTGGCACTTCAGTGGCACTTCAGTAAACACATGACCGG 3299
Qy 301 GACGACACTATGCTGCTGCTGCTTTGTTGTTACAGTCAAGCAGCCGCTTCTATGCTGTG 360
Db 3300 GACGACACTATGCTGCTGCTGCTTTGTTGTTACAGTCAAGCAGCCGCTTCTATGCTGTG 3359
Qy 361 ATGTGGAAGCAGGTGACGACAGCTACTTGGGAGACCAAGCCACCGCGGCTTATGGCTAC 420
Db 3360 ATGTGGAAGCAGGTGACGACAGCTACTTGGGAGACCAAGCCACCGCGGCTTATGGCTAC 3419
Qy 421 TCCGGCGTGTCCCTCAAGGTGGTGAATCTCCACAGGGGACGGCGAGCACCTCAGGAAC 480
Db 3420 TCCGGCGTGTCCCTCAAGGTGGTGAATCTCCACAGGGGACGGCGAGCACCTCAGGAAC 3479
Qy 481 GCGCTGTGGCACAGGGGACACCGCGGGGAGGTGGAACTTTATGACAGACCCGAGG 540
Db 3480 GCGCTGTGGCACAGGGGACACCGCGGGGAGGTGGAACTTTATGACAGACCCGAGG 3539
Qy 541 AACATTGGCTGGAAGGACTACACGGCTCTATAGGTGGCACCTGACTCAGAGCCCAAGACC 600
Db 3540 AACATTGGCTGGAAGGACTACACGGCTCTATAGGTGGCACCTGACTCAGAGCCCAAGACC 3599
Qy 601 GGCTACATCAGAGCTTTAGTGCATGAAGGAAAAAGGTCTATGAGTGGCACTCAGGACCTATC 660
Db 3600 GGCTACATCAGAGCTTTAGTGCATGAAGGAAAAAGGTCTATGAGTGGCACTCAGGACCTATC 3659
Qy 661 TATGACCAAACTACGCTGGGGGGGCGTGGTCTATTTGTTCTCTCAAGAAATGTC 720
Db 3660 TATGACCAAACTACGCTGGGGGGGCGTGGTCTATTTGTTCTCTCAAGAAATGTC 3719
Qy 721 TATTTCTCAGACCTCAAGTACGAATGACAGAGATATTTAAACAAGATTTGCTGCAATTCG 780
Db 3720 TATTTCTCAGACCTCAAGTACGAATGACAGAGATATTTAAACAAGATTTGCTGCAATTCG 3779
Qy 781 GCAATGCCCTGTCATGTCATGTCCTCCTAGACACCTCAGTTCATTTGCTGCTCGGCT 840
Db 3780 GCAATGCCCTGTCATGTCATGTCCTCCTAGACACCTCAGTTCATTTGCTGCTCGGCT 3839
Qy 841 TCTCTCTAGCAGACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 3840 TCTCTCTAGCAGACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3899
Qy 901 CCAGCAACCCCAACCCCAAGTCCCTTCAGAGGATAAATATCAATGGAATCTCAGAGTAA 960
Db 3900 CCAGCAACCCCAACCCCAAGTCCCTTCAGAGGATAAATATCAATGGAATCTCAGAGTAA 3959
Qy 961 CATCTAAACCACTAGAGGAAACAGTTGGTGTATATAGAGCTTTATGTGGAGTGAAA 1020
Db 3960 CATCTAAACCACTAGAGGAAACAGTTGGTGTATATAGAGCTTTATGTGGAGTGAAA 4019
Qy 1021 TTGGGATCCCAATCATTTGCTTTTCTGTTTGTGTTTAAAGAAATGAGCTTTACATATA 1080
Db 4020 TTGGGATCCCAATCATTTGCTTTTCTGTTTGTGTTTAAAGAAATGAGCTTTACATATA 4079
Qy 1081 AAATGTAAATTAATTTATTTATGTTATGTTATGAGTTGAAAGGGAATCTGTGCATAAG 1140
Db 4080 AAATGTAAATTAATTTATTTATGTTATGTTATGAGTTGAAAGGGAATCTGTGCATAAG 4139
Qy 1141 CCATTTATGATAAATTAAGCATGAATAATTTGCTGAACTACTTTGGCTTAAAGTTCT 1200
Db 4140 CCATTTATGATAAATTAAGCATGAATAATTTGCTGAACTACTTTGGCTTAAAGTTCT 4199
Qy 1201 CACTATTTGTAATTAGAGTTGCTCTACAATGACACAAATCCCGCTAAATAAATATA 1260
Db 4200 CACTATTTGTAATTAGAGTTGCTCTACAATGACACAAATCCCGCTAAATAAATATA 4259
Qy 1261 AACAGGGTCAATTCAAATTTGAAGTAAATTTTATGAGGAGATTTAGAGACCAACG 1320

Db 4260 AACAGGGTCAATTCAAATTTGAAGTAAATTTTATGAGGAGATTTAGAGACAACAG 4319
Qy 1321 GCATAGCAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAACAGTTTACAAA 1380
Db 4320 GCATAGCAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAACAGTTTACAAA 4379
Qy 1381 ATAAACGAACCT 1440
Db 4380 ATAAACGAACCT 4439
Qy 1441 TCATCAAGAACCAACATCT 1500
Db 4440 TCATCAAGAACCAACATCT 4499
Qy 1501 AAAACCTCATTTAAGTAAAAAGTTAGCAGACAAAGTGGCGTCTCTCTCTCTCTCTCTCTCT 1560
Db 4500 AAAACCTCATTTAAGTAAAAAGTTAGCAGACAAAGTGGCGTCTCTCTCTCTCTCTCTCT 4559
Qy 1561 GCCGTTGGGCTCGGGAGGCT 1620
Db 4560 GCCGTTGGGCTCGGGAGGCT 4619
Qy 1621 AGGAACAGAGCAGACGACGAGCCGGAAGGGCGCATCTAAACGCTATCTAGGCTTTGG 1680
Db 4620 AGGAACAGAGCAGACGACGAGCCGGAAGGGCGCATCTAAACGCTATCTAGGCTTTGG 4679
Qy 1681 TAACTGGGCAAGTCT 1740
Db 4680 TAACTGGGCAAGTCT 4739
Qy 1741 TAAATATTTGTTAATATTTAATTAAGTACATATAGATGATACATTTTCAATTAAGTTTCAAGTTA 1800
Db 4740 TAAATATTTGTTAATATTTAATTAAGTACATATAGATGATACATTTTCAATTAAGTTTCAAGTTA 4799
Qy 1801 TATTTTAAATGCTCTAGTACACATATAGTATATATTTCTAGAAAACAAACATCTAATA 1860
Db 4800 TATTTTAAATGCTCTAGTACACATATAGTATATATTTCTAGAAAACAAACATCTAATA 4859
Qy 1861 AGTATATACTCTGTGAAAATATGAGGCTTGATATATATTTAGTTGTCTGATGAAGATG 1920
Db 4860 AGTATATACTCTGTGAAAATATGAGGCTTGATATATATTTAGTTGTCTGATGAAGATG 4919
Qy 1921 CTAGAGCTGTAAACAGAAATACATAGAGAAATATGAGGAGTTTATGATGGAACCTTAAAT 1980
Db 4920 CTAGAGCTGTAAACAGAAATACATAGAGAAATATGAGGAGTTTATGATGGAACCTTAAAT 4979
Qy 1981 ATATGTTGCCAGCGATTTTATTTCAATATTTGTTTACTGTTTATCTATCTGCTGTATG 2040
Db 4980 ATATGTTGCCAGCGATTTTATTTCAATATTTGTTTACTGTTTATCTATCTGCTGTATG 5039
Qy 2041 AATCTTTTAAATTCAAACGCTGAAAACGAATCAGCATTTAGTCTTGGCAGGACACCCAA 2100
Db 5040 AATCTTTTAAATTCAAACGCTGAAAACGAATCAGCATTTAGTCTTGGCAGGACACCCAA 5099
Qy 2101 TAACTAGTCTGTAAATATGCAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2160
Db 5100 TAACTAGTCTGTAAATATGCAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 5159
Qy 2161 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2220
Db 5160 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 5219
Qy 2221 TAAGGGTTTGTAGTAAAGAGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2280
Db 5220 TAAGGGTTTGTAGTAAAGAGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 5279
Qy 2281 CCGTTCTGTTTAAATAGTCTATCTATCTATCCCAACAGGAACGATCCATCACTTACTT 2340
Db 5280 CCGTTCTGTTTAAATAGTCTATCTATCTATCCCAACAGGAACGATCCATCACTTACTT 5339
Qy 2341 CTTAATGTACATTTGCAATTTTGAATAAATTAATTTGTTTGTGTTTGTGTTTGTGTTTGTG 2400
Db 5340 CTTAATGTACATTTGCAATTTTGAATAAATTAATTTGTTTGTGTTTGTGTTTGTGTTTGTG 5399

```
QY 2401 TTGTTGTTGTTGTCGACATTTTACTTTTTCGTTGGAGCTGTATTCGCCGAGCAAC 2460
Db 5400 TTGTTGTTGTTGTCGACATTTTACTTTTTCGTTGGAGCTGTATTCGCCGAGCAAC 5459
QY 2461 GAAGCGTTGGGATCTTCAATTAATGAGGAGCTGTCAACAGCGTGTTCCTGTTT 2520
Db 5460 GAAGCGTTGGGATCTTCAATTAATGAGGAGCTGTCAACAGCGTGTTCCTGTTT 5519
QY 2521 CTGTTGTTGGGCTCAACCGTACAAATGTTGGGATGAGGAGTGAATATTTAGAA 2580
Db 5520 CTGTTGTTGGGCTCAACCGTACAAATGTTGGGATGAGGAGTGAATATTTAGAA 5579
QY 2581 TGTACCAATATTTTCTAAATATTTATGTTTCTTAAACAAATTTATCGTAGTTTGA 2640
Db 5580 TGTACCAATATTTTCTAAATATTTATGTTTCTTAAACAAATTTATCGTAGTTTGA 5639
QY 2641 TGAAGGCTATGTTTTCGCAAGAGCTGTAAATATTTATTAAGTTTCAATGTCAA 2700
Db 5640 TGAAGGCTATGTTTTCGCAAGAGCTGTAAATATTTATTAAGTTTCAATGTCAA 5699
QY 2701 AATTTCAACACTGAAACCTGACCTTAGCTAGAACCTCATTTTAAAGATTAACACAG 2760
Db 5700 AATTTCAACACTGAAACCTGACCTTAGCTAGAACCTCATTTTAAAGATTAACACAG 5759
QY 2761 AATAAATTTGAAAAAGGTTTCT 2785
Db 5760 AATAAATTTGAAAAAGGTTTCT 5784

RESULT 2
AX375587
LOCUS AX375587 5784 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 3 from Patent WO0209735.
ACCESSION AX375587
VERSION AX375587.1 GI:19170155
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Bornstein, P., Kyriakides, T., Ratner, B., Giachelli, C.,
Martinson, L.D. and Scatena, M.
Methods and devices to modulate the wound response
Patent: WO 0209735-A 3 07-FEB-2002;
The University of Washington (US)
FEATURES
Location/Qualifiers
1..5784
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
240..3758
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD26800.1"
/db_xref="GI:19170155"
/translation="MVMRLVLLALWVWPSTQAGHQDKDTPDLFSISINIRKTTGAKQ
FRGPDGPVAFVRFVDYIPVNADDLSKITKMRQKGFELTALQKQDGKSRGLLA
YSLHGLQSRQREIVSNGPADFLDLYMIDGTRHVVSLEVDGLADSKQWNTVQVAGET
YSLHGLQSRQREIVSNGPADFLDLYMIDGTRHVVSLEVDGLADSKQWNTVQVAGET
DISSKQSQQQQGAIEINAISENTELRGLPHVTTEYVGPSSERRPEVCERSELGNM
VOELSGHLVNLQSLSENLRVNSNDNQLWELIGGPPKTRNMSACWQGRFPAENETW
VDSCTCTCKKFKFTCHQITCPATCASPSFVEGECPCSLHSDVBEGHSPNABWTQ
CSVTCCGSGTOQRGSCDVTNCLGPSIQTRACSLSKDTRIRQDGGWSHSPWSSCS
VTCGVGNITRILNCSVPVPMGKNGKSGRETKACQAPCIDGRWSHSPWSSCSV
TCAGIRERTKVCNPSFPQYGGKACVGDVQRCNKRSCFVDCVCLNPPCPGACQSS
FPDGSWSGCFPGVLNGTHCEDLDECALVPDI CFSTSKVPRCWNTQPGHPLCPPP
RYRGNQPVGLEDLAKEKQVCEPNCKTKHCHAEICILYGFHSDPMYKCECOT
GYAGDGLICGSDLDLGNPILNLCATNATVHCIKDNCPLHNSGDEDFKGLGDAC
DDDDNDGVDTEKDNCLLFPNRQADYDKVEGVRCDNCPIYHNPQITDITDNNEGDA
CSVDIDGDDVFNERNKCPYVNTQORDTDGVDGVDHCDNCPVHNPQITDITDNNELVG
QCDNNEDIDDDGHQNNQDNCFYIISNANQAHDHDDGQGDACDPDDDDNDGVDPRDNCRL
```

```
VENPOEDLDGDRGDIKCDKDFDNDONIPDIDDDVCPENNAISEITDFRNFQWFLDPKGT
TQIDNNWIRHOGKELNQTANSDPGIAGVDFSGTFTVYNTDRDDYAGVFG
YSSSRFYVMKQVTQTYWEDQPTRAYGVSGLKVNSTGTGTGHLRNALWHTGNT
PGQVTLRHDPRNIGWKDYTAIRWHLTHRPKTYIRLVLHVEGKQVMADSGPIYDQTYA
GGRLGLFVPSQBMVYFSDLKYBCRDI"
ORIGIN
Query Match 100.0%; Score 2785; DB 6; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTCGACGGTGATCGACGGGTGATATTTGTAAGATGATTTTGACAAATGACAAC 60
Db 3000 GAGGACTTCGACGGTGATCGACGGGTGATATTTGTAAGATGATTTTGACAAATGACAAC 3059
QY 61 ATCCAGATATTTGATGATGTGTCTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 120
Db 3060 ATCCAGATATTTGATGATGTGTCTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 3119
QY 121 AACTTCCAGATGTTCCCTTGGATCCAAAGGAGACACCCAAATTTGATCCCACTGGGTC 180
Db 3120 AACTTCCAGATGTTCCCTTGGATCCCAAGGAGACACCCAAATTTGATCCCACTGGGTC 3179
QY 181 ATTCCGCATCAAGGCAAGGAGCTGTTCAGACAGCAACTCGGACCCCGGCATCGCTGA 240
Db 3180 ATTCCGCATCAAGGCAAGGAGCTGTTCAGACAGCAACTCGGACCCCGGCATCGCTGA 3239
QY 241 GGTTTGACAGTGTGGGTCTGTGACATTCAGTGGCACATTTACGTAAACACTGACCCG 300
Db 3240 GGTTTGACAGTGTGGGTCTGTGACATTCAGTGGCACATTTACGTAAACACTGACCCG 3299
QY 301 GACGAGCACTATGCTGGCTTCTGTTTGGTTACCAAGTCAAGCAGCGCTTCTATGTTGGT 360
Db 3300 GACGAGCACTATGCTGGCTTCTGTTTGGTTACCAAGTCAAGCAGCGCTTCTATGTTGGT 3359
QY 361 ATGTGGAAGCAGGTGACGACGACCTTACTGGGAGGACACCCACCGGCGGCTTATGGCTAC 420
Db 3360 ATGTGGAAGCAGGTGACGACGACCTTACTGGGAGGACACCCACCGGCGGCTTATGGCTAC 3419
QY 421 TCCGCGGTGTCCTCAAGGTGGTGAATCCACACGGGAGCGGAGCAGCCTTAGAGAAC 480
Db 3420 TCCGCGGTGTCCTCAAGGTGGTGAATCCACACGGGAGCGGAGCAGCCTTAGAGAAC 3479
QY 481 GCGCTGTGGCACAGGGGAAACAGCCGGGGCAGGTCGAACTTATGACAGACCCAGG 540
Db 3480 GCGCTGTGGCACAGGGGAAACAGCCGGGGCAGGTCGAACTTATGACAGACCCAGG 3539
QY 541 AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGCACCTGACTCAGGCCCCAAGACC 600
Db 3540 AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGCACCTGACTCAGGCCCCAAGACC 3599
QY 601 GGCTACATCAGAGCTTTAGTGCATGAAGGAAAAACAGGTTCATGGCAGACTCAGGACTATC 660
Db 3600 GGCTACATCAGAGCTTTAGTGCATGAAGGAAAAACAGGTTCATGGCAGACTCAGGACTATC 3659
QY 661 TATGACCAAACTACGCTGGCGGGGCTGGGTCTATTTGTTCTTCTCAAGAAATGTC 720
Db 3660 TATGACCAAACTACGCTGGCGGGGCTGGGTCTATTTGTTCTTCTCAAGAAATGTC 3719
QY 721 TATTTCTCAGACCTCAAGTACGAATGAGAGATATTTAAACAAGATTTGCTGCAATTTCCG 780
Db 3720 TATTTCTCAGACCTCAAGTACGAATGAGAGATATTTAAACAAGATTTGCTGCAATTTCCG 3779
QY 781 GCAATGCCCTGTGCATGCCATGGTCCCTAGACACTCAGTTTCTTGTGCTTCTTGGCGCT 840
Db 3780 GCAATGCCCTGTGCATGCCATGGTCCCTAGACACTCAGTTTCTTGTGCTTCTTGGCGCT 3839
QY 841 TCTCTCTTAGCAGCACTCTGTCCTCTGACCTTAACCTGATGTTCTTCTCACTCTG 900
Db 3840 TCTCTCTTAGCAGCACTCTGTCCTCTGACCTTAACCTGATGTTCTTCTCACTCTG 3899
QY 901 CCAGCAACCCCAACCCCAAGTCCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960
Db 901 CCAGCAACCCCAACCCCAAGTCCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960
```


ORIGIN		/mol_type="unassigned DNA" /db_xref="taxon:9606" /notes="EMBL/GenBank Accession No. L12350"	
Query Match		100.0%;	Score 2785; DB 6; Length 5784;
Best Local Similarity		100.0%;	Pred. No. 0;
Matches 2785; Conservative 0;		Mismatches 0;	Indels 0; Gaps 0;
QY	1	GAGGACTTGGACCGGTGATGACGGGGTGTATTTGTAAGATCATTTTGACATGCAAC	60
DB	3000	GAGGACTTGGACCGGTGATGACGGGGTGTATTTGTAAGATCATTTTGACATGCAAC	3059
QY	61	ATCCAGATATTGATGATGTGTCTCTGAAAAAACAATGCCATCATGTGAGACAGACTTCAGG	120
DB	3060	ATCCAGATATTGATGATGTGTCTCTGAAAAAACAATGCCATCATGTGAGACAGACTTCAGG	3119
QY	121	AACCTCCAGATGTCCTTGGATCCCAAGGACCAACCCAAATGATCCCAACTGGGTC	180
DB	3120	AACCTCCAGATGTCCTTGGATCCCAAGGACCAACCCAAATGATCCCAACTGGGTC	3179
QY	181	ATTTCGCATCAAGGCAAGAGCTGTGTTACAGACGCCAACTCGGACCCCGGCATCGCTGTA	240
DB	3180	ATTTCGCATCAAGGCAAGAGCTGTGTTACAGACGCCAACTCGGACCCCGGCATCGCTGTA	3239
QY	241	GGTTTGTGACGATTTGGGTCTGTGGACTTCAGTGGCACATTTCTACGTAAACACTGACCGG	300
DB	3240	GGTTTGTGACGATTTGGGTCTGTGGACTTCAGTGGCACATTTCTACGTAAACACTGACCGG	3299
QY	301	GACGACGATGCTGGCTTCCTTGTGTACAGTCAAGGACCGGCTTCATGTGGTG	360
DB	3300	GACGACGATGCTGGCTTCCTTGTGTACAGTCAAGGACCGGCTTCATGTGGTG	3359
QY	361	ATGTGGAAGCAGGTGACGACGACCTACTCGGAGGACCGGAGCACCTTACGTAAGACCTGAGGTA	420
DB	3360	ATGTGGAAGCAGGTGACGACGACCTACTCGGAGGACCGGAGCACCTTACGTAAGACCTGAGGTA	3419
QY	421	TCCGGGTGTCCTCAAGGTGTGAACTCCACGCGGAGCGGGAGCACCTTACGTAAGAC	480
DB	3420	TCCGGGTGTCCTCAAGGTGTGAACTCCACGCGGAGCGGGAGCACCTTACGTAAGAC	3479
QY	481	GGCTGTGGCAACCGGGNACACGCGGGGAGGTGCGAACCTTATGGACACCCACG	540
DB	3480	GGCTGTGGCAACCGGGNACACGCGGGGAGGTGCGAACCTTATGGACACCCACG	3539
QY	541	AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGCACTTACGACGCGGCGGAGACC	600
DB	3540	AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGCACTTACGACGCGGCGGAGACC	3599
QY	601	GGCTACATCAGAGTCTTAGTGATCAAGGAAAAACAGGTCTAGGACCTCAGGACCTATC	660
DB	3600	GGCTACATCAGAGTCTTAGTGATCAAGGAAAAACAGGTCTAGGACCTCAGGACCTATC	3659
QY	661	TATGACCAAACTACGCTGGGGGGGCTGGGTCTATTGTCTCTCTCAAGAAATGTC	720
DB	3660	TATGACCAAACTACGCTGGGGGGGCTGGGTCTATTGTCTCTCTCAAGAAATGTC	3719
QY	721	TATTTCTCAGACCTCAAGTACGAATGACAGATATTTTAAACAAGATTTGCTGCAATTCGG	780
DB	3720	TATTTCTCAGACCTCAAGTACGAATGACAGATATTTTAAACAAGATTTGCTGCAATTCGG	3779
QY	781	GCAATGCCCTGTGCATGCCATGGTCCCTTAGACACCTCAGTTCAATGTGGTCTCTGCGGCT	840
DB	3780	GCAATGCCCTGTGCATGCCATGGTCCCTTAGACACCTCAGTTCAATGTGGTCTCTGCGGCT	3839
QY	841	TCTCTCTACGACGACCTCTGTCCCTTACCTTAACTCTGATGGTTCCTTCACTCTCG	900
DB	3840	TCTCTCTACGACGACCTCTGTCCCTTACCTTAACTCTGATGGTTCCTTCACTCTCG	3899
QY	901	CCAGCAACCCCAACCCCAAGTCCCTTACAGAGATATAATCAATGGAACCTCAGAGATGA	960
DB	3900	CCAGCAACCCCAACCCCAAGTCCCTTACAGAGATATAATCAATGGAACCTCAGAGATGA	3959

QY	961	CATCTAACCCACTAGAGGAAACCAGTTTGGTGATATATGAGACTTTATGTGGAGTGAAAA	1020
DB	3960	CATCTAACCCACTAGAGGAAACCAGTTTGGTGATATATGAGACTTTATGTGGAGTGAAAA	4019
QY	1021	TTGGGCATGCCAATTAATTCCTTTTCTTTTGTGTTTAAAAAGAAATGACCTTTACATATA	1080
DB	4020	TTGGGCATGCCAATTAATTCCTTTTCTTTTGTGTTTAAAAAGAAATGACCTTTACATATA	4079
QY	1081	AAATGTAAATTAATTAATTTATGTGTATATTTGAGGTGAAAGGAATCTGTGCATAAG	1140
DB	4080	AAATGTAAATTAATTAATTTATGTGTATATTTGAGGTGAAAGGAATCTGTGCATAAG	4139
QY	1141	CAATTATGATAAATTAAGCATGAATAATTCCTGAACCTTTTGGTGTCTTAAAGTTGT	1200
DB	4140	CAATTATGATAAATTAAGCATGAATAATTTGCTGAACCTTTTGGTGTCTTAAAGTTGT	4199
QY	1201	CACATTTCTTGAATTTAGAGTTGCTCTACAAATGACACACAAATCCCGCTAAATAAATATA	1260
DB	4200	CACATTTCTTGAATTTAGAGTTGCTCTACAAATGACACACAAATCCCGCTAAATAAATATA	4259
QY	1261	AACAAGGTCATTTCAATTTGAAATTTAGTAAAGGAGAGATTAGAAGACAACAG	1320
DB	4260	AACAAGGTCATTTCAATTTGAAATTTAGTAAAGGAGAGATTAGAAGACAACAG	4319
QY	1321	GCATAGCAAAATGACATTAAGCTACCGATTAACTAAATCGGAACATGTAAACAGTTACAAA	1380
DB	4320	GCATAGCAAAATGACATTAAGCTACCGATTAACTAAATCGGAACATGTAAACAGTTACAAA	4379
QY	1381	ATAAACGAACTCTCCTTCTTCTTCAATTAAGGAGGCTCATGTGACAGTAGAGATGAGTT	1440
DB	4380	ATAAACGAACTCTCCTTCTTCTTCAATTAAGGAGGCTCATGTGACAGTAGAGATGAGTT	4439
QY	1441	TCATCAAAAGAAACAAATCTTTGAAATTTGAGGTGTGACGCGTTCCAGATGTGATTTGGC	1500
DB	4440	TCATCAAAAGAAACAAATCTTTGAAATTTGAGGTGTGACGCGTTCCAGATGTGATTTGGC	4499
QY	1501	AAAACTCATTTTAAAGTTTACGAGCAAGAGTGGGTGCTTTAGCTGCTGCTTGT	1560
DB	4500	AAAACTCATTTTAAAGTTTACGAGCAAGAGTGGGTGCTTTAGCTGCTGCTTGT	4559
QY	1561	GCGGTGTGGGTGCGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG	1620
DB	4560	GCGGTGTGGGTGCGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG	4619
QY	1621	AGGAACAGAGCAGACGACAGCGCGGAAAGGCGCATCTAACGCGTATCTAGGCTTTGG	1680
DB	4620	AGGAACAGAGCAGACGACAGCGCGGAAAGGCGCATCTAACGCGTATCTAGGCTTTGG	4679
QY	1681	TAACTGCGGACAGGTTGCTTTTACCTGATTTTGTATGATACATTTCAATTAAGGTTCCAGTTA	1740
DB	4680	TAACTGCGGACAGGTTGCTTTTACCTGATTTTGTATGATACATTTCAATTAAGGTTCCAGTTA	4739
QY	1741	TAAATATTTTGTAAATATTTTAAAGTACATATAGAAATGCAACTCCATTTACCAGTAAC	1800
DB	4740	TAAATATTTTGTAAATATTTTAAAGTACATATAGAAATGCAACTCCATTTACCAGTAAC	4799
QY	1801	TATTTTAAATATGCTAGTAAACATATAGTATAATTTCTAGAAAACAAACATCTAATA	1860
DB	4800	TATTTTAAATATGCTAGTAAACATATAGTATAATTTCTAGAAAACAAACATCTAATA	4859
QY	1861	AGTATATAATCTCTGAAATATGAGGCTTGATATAATTTAGGTTGTACAGTGAAGCATG	1920
DB	4860	AGTATATAATCTCTGAAATATGAGGCTTGATATAATTTAGGTTGTACAGTGAAGCATG	4919
QY	1921	CTAGAAGCTGTAAACAGAAATACATAGAGAAATAAGGAGTTTATGATGGACCTTAATAT	1980
DB	4920	CTAGAAGCTGTAAACAGAAATACATAGAGAAATAAGGAGTTTATGATGGACCTTAATAT	4979
QY	1981	ATAATGTTCAGCGATTTTGTAGTTCAATATTTGTGTACTGTATCTATCTGCTGTATATGG	2040
DB	4980	ATAATGTTCAGCGATTTTGTAGTTCAATATTTGTGTACTGTATCTATCTGCTGTATATGG	5039
QY	2041	AAATCTTTTAAATCAACGCTGAAACAGAAATCAGATTTTGTCTTTGCCAGCACCCCAA	2100

3960 CATCTAACCCCTAGAGGAAACCAGTTTGGTGATATAGAGACTTTATGTGGAGTGA AAA 4019
1021 TTGGGCATGCCATTACATTTGCTTTTCTGTTGTTTAAAGAAAGAGCTTTACATATA 1080
4020 TTGGGCATGCCATTACATTTGCTTTTCTGTTTAAAGAAAGAGCTTTACATATA 4079
1081 AAATGTAATTA CTTATTGTTATTATGTTGATATGAGTTGAAAGGGAATACCTGTCATAAG 1140
4080 AAATGTAATTA CTTATTGTTATTATGTTGATATGAGTTGAAAGGGAATACCTGTCATAAG 4139
1141 CCAATTATGTAATTAATTAAGCATGAATAATATGCTGAACTACTCTTTTGGTCTTAAAGTTGT 1200
4140 CCAATTATGTAATTAATTAAGCATGAATAATATGCTGAACTACTCTTTTGGTCTTAAAGTTGT 4199
1201 CACTATTCTTGAATTAAGTGTGCTCTACAAATGACACACAAATCCCGCTAAATAAATTATA 1260
4200 CACTATTCTTGAATTAAGTGTGCTCTACAAATGACACACAAATCCCGCTAAATAAATTATA 4259
1261 AACAAAGGTCATTTCAAATTTGAAAGTAATGTTTATGTAAGGAGAGATTAGAAAGACAAACAG 1320
4260 AACAAAGGTCATTTCAAATTTGAAAGTAATGTTTATGTAAGGAGAGATTAGAAAGACAAACAG 4319
1321 GCATAGCAAAATGACATAGCTACCGATTAACCTAATCGGAAACATGTAATAACAGTTTCAAAA 1380
4320 GCATAGCAAAATGACATAGCTACCGATTAACCTAATCGGAAACATGTAATAACAGTTTCAAAA 4379
1381 ATAAACGAACTCTCTCTGCTCTCAATGAAAGCCCTCATGTGAGTAGAGATCGAGTT 1440
4380 ATAAACGAACTCTCTCTGCTCTCAATGAAAGCCCTCATGTGAGTAGAGATCGAGTT 4439
1441 TCATCAAAAGAAACAAACATCTTCTGCAATGAGGTGTGACGCGTTCCAGATGTGGATTGGC 1500
4440 TCATCAAAAGAAACAAACATCTTCTGCAATGAGGTGTGACGCGTTCCAGATGTGGATTGGC 4499
1501 AAAACCTCATTTAAGTAAAGGTTAGCAGAGCAAAAGTGGCTTTAGCTGCTGCTGTGT 1560
4500 AAAACCTCATTTAAGTAAAGGTTAGCAGAGCAAAAGTGGCTTTAGCTGCTGCTGTGT 4559
1561 GCCGTTGTGGGCTCGGGAGGCTTCTGCTGAGCTTCTCTCCAGCTTTGCTGCTGCTGAG 1620
4560 GCCGTTGTGGGCTCGGGAGGCTTCTGCTGAGCTTCTCTCCAGCTTTGCTGCTGCTGAG 4619
1621 AGGAACCAAGAGCAGCAGCAGCGGAAAGGCGCATCTACGCGTATCTAGGCTTTGG 1680
4620 AGGAACCAAGAGCAGCAGCAGCGGAAAGGCGCATCTACGCGTATCTAGGCTTTGG 4679
1681 TAACGCGGACAGAGTTGCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA 1740
4680 TAACGCGGACAGAGTTGCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA 4739
1741 TAAATATTTTGTAAATATTTAATGAGTGACATAGAAATGCACTCCATTTACCAGTAACT 1800
4740 TAAATATTTTGTAAATATTTAATGAGTGACATAGAAATGCACTCCATTTACCAGTAACT 4799
1801 TATTTTAAATAGCTAGTAACACATATGATGATATAATTTCTAGAAACAAACATCTAATA 1860
4800 TATTTTAAATAGCTAGTAACACATATGATGATATAATTTCTAGAAACAAACATCTAATA 4859
1861 AGTATATAATCTGTGAAATATGAGGCTTGATAATATTAAGTTGTCGATGAAGCATG 1920
4860 AGTATATAATCTGTGAAATATGAGGCTTGATAATATTAAGTTGTCGATGAAGCATG 4919
1921 CTAGAAGCTGTAAACAGATACATAGAAATATGAGGAGTTTATGATGAACCTTAAATAT 1980
4920 CTAGAAGCTGTAAACAGATACATAGAAATATGAGGAGTTTATGATGAACCTTAAATAT 4979
1981 ATAATGTTGCCAGGATTTTGTGTTCAATATTTGTTTACTGTTATCTATCTGCTGATATGG 2040
4980 ATAATGTTGCCAGGATTTTGTGTTCAATATTTGTTTACTGTTATCTATCTGCTGATATGG 5039
2041 AATTCCTTTTAAATTAAGCTGAAACGAATCAGCATTTAGCTTTGCCAGGACACCCAA 2100
5040 AATTCCTTTTAAATTAAGCTGAAACGAATCAGCATTTAGCTTTGCCAGGACACCCAA 5099

QY 2101 TAATCAGTCATGTGTAAATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2160
Db 5100 TAATCAGTCATGTGTAAATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 5159
QY 2161 TTTTGTGCTTTAAGTTGTCATGATCTTCTGAGGAAATAGTCACTCATCCCACTCCACA 2220
Db 5160 TTTTGTGCTTTAAGTTGTCATGATCTTCTGAGGAAATAGTCACTCATCCCACTCCACA 5219
QY 2221 TAAGGGTTTACTAAGAGAAAGTCTGTCTGTCTGATGATGATAGGGGGCAATCTTTTTC 2280
Db 5220 TAAGGGTTTACTAAGAGAAAGTCTGTCTGTCTGATGATGATAGGGGGCAATCTTTTTC 5279
QY 2281 CCGTTTCTGTTAATAGTCACTCATCTATGCGAAACAGGAACGATCCATTAACCTTAGT 2340
Db 5280 CCGTTTCTGTTAATAGTCACTCATCTATGCGAAACAGGAACGATCCATTAACCTTAGT 5339
QY 2341 CTTAATGTACACATTTGCACTTTTGTAAATAATTTGTTGTTTCTCTTGGAGTTGATCG 2400
Db 5340 CTTAATGTACACATTTGCACTTTTGTAAATAATTTGTTGTTTCTCTTGGAGTTGATCG 5399
QY 2401 TTGTGTTTGTGTCACCTTTTACTTTTTCGTTGTGGAGCTGTATTTCCCGAGACAAC 2460
Db 5400 TTGTGTTTGTGTCACCTTTTACTTTTTCGTTGTGGAGCTGTATTTCCCGAGACAAC 5459
QY 2461 GAAGCGTTGGGATCTTCAATTAATGTAGCAGCTGTCAACAGCGTGCAGGTTTCTGTTT 2520
Db 5460 GAAGCGTTGGGATCTTCAATTAATGTAGCAGCTGTCAACAGCGTGCAGGTTTCTGTTT 5519
QY 2521 CTGCTTCTGGGGTCAACCGTACAATGCTGGGAATGACGATGATGGAATATTAGAA 2580
Db 5520 CTGCTTCTGGGGTCAACCGTACAATGCTGGGAATGACGATGATGGAATATTAGAA 5579
QY 2581 TGTACCAATATTTTGTAAATATTTATGTTTCTTAAACAAATTTATCGTATAGTTGA 2640
Db 5580 TGTACCAATATTTTGTAAATATTTATGTTTCTTAAACAAATTTATCGTATAGTTGA 5639
QY 2641 TGAAACGTCATGTTTCCCAAGACCTGTAAATATTTATGTTTATGTTTACATGTTCA 2700
Db 5640 TGAAACGTCATGTTTCCCAAGACCTGTAAATATTTATGTTTATGTTTACATGTTCA 5699
QY 2701 AATTTCCACCTGAAACCCCTGACCTTAGCACTCATTTTAAAGATTAAACACAGG 2760
Db 5700 AATTTCCACCTGAAACCCCTGACCTTAGCACTCATTTTAAAGATTAAACACAGG 2760
QY 2761 AAATAAATTTGTAAGGTTTCT 2785
Db 5760 AAATAAATTTGTAAGGTTTCT 5784

RESULT 5
AX774982
LOCUS AX774982 5784 bp DNA linear PAT 09-JUL-2003
DEFINITION Sequence 298 from Patent WO03038129.
ACCESSION AX774982
VERSION AX774982.1 GI:32486498
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Raponi, M.
METHODS Methods for assessing and treating leukemia
Patent: WO 03038129-A 298 08-MAY-2003;
Ortho-Clinical Diagnostics, Inc. (US)
FEATURES
Location/Qualifiers
1..5784
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 2785; DB 6; Length 5784; Best Local Similarity 100.0%; Pred. No. 0; Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	GAGGACTTGGACGGTGTATGACGGGGTGAATTTGTTAAAGATGATTTTGGACAATGACAAC	60						
Db	3000	GAGGACTTGGACGGTGTATGACGGGGTGAATTTGTTAAAGATGATTTTGGACAATGACAAC	3059						
Qy	61	ATCCAGATATTGATGATGTGTCTGCTGAAACCAATGCTCAGTGAGACAGACTTCAGG	120						
Db	3060	ATCCAGATATTGATGATGTGTCTGCTGAAACCAATGCTCAGTGAGACAGACTTCAGG	3119						
Qy	121	AACCTTCCAGATGTCCCTTGGATCCCAAGGGGACCCAAATTTGATCCCACTGGGTC	180						
Db	3120	AACCTTCCAGATGTCCCTTGGATCCCAAGGGGACCCAAATTTGATCCCACTGGGTC	3179						
Qy	181	ATTGCCCATCAAGGCAAGGAGCTGGTTGACAGCAGCACTCGGACCCCGGATCGCTGTA	240						
Db	3180	ATTGCCCATCAAGGCAAGGAGCTGGTTGACAGCAGCACTCGGACCCCGGATCGCTGTA	3239						
Qy	241	GGTTTTCAGAGTTTGGTCTGTGAGCTTTCAGTGGCACAATTCAGTAAACACTGACCGG	300						
Db	3240	GGTTTTCAGAGTTTGGTCTGTGAGCTTTCAGTGGCACAATTCAGTAAACACTGACCGG	3299						
Qy	301	GACGACGACTATGCTGGCTTCCTTTTGGTTTACCAAGTCAAGCAGCCGCTTCTATGTGGTG	360						
Db	3300	GACGACGACTATGCTGGCTTCCTTTTGGTTTACCAAGTCAAGCAGCCGCTTCTATGTGGTG	3359						
Qy	361	ATGTGGAAAGCAGGTGACGACACCTTAATGAGGAGACAGCCCAACGCGGGCTTATGGCTAC	420						
Db	3360	ATGTGGAAAGCAGGTGACGACACCTTAATGAGGAGACAGCCCAACGCGGGCTTATGGCTAC	3419						
Qy	421	TCCGCGGTGCTCCCTCAAGGTGTGAGTTCACCAAGGGGAGCGGGGAGCAGCTCAGAGAAC	480						
Db	3420	TCCGCGGTGCTCCCTCAAGGTGTGAGTTCACCAAGGGGAGCGGGGAGCAGCTCAGAGAAC	3479						
Qy	481	GCGCTGTGGCACAACGGGGAACACGCGGGGAGGTGCGAACCTTATGACGACGCCACGG	540						
Db	3480	GCGCTGTGGCACAACGGGGAACACGCGGGGAGGTGCGAACCTTATGACGACGCCACGG	3539						
Qy	541	AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGGACCTGACTCAGAGCCCAAGACC	600						
Db	3540	AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGGACCTGACTCAGAGCCCAAGACC	3599						
Qy	601	GGCTACATCAGAGTCTTAGTGATGAAGAAACAGGTATGCGGACAGCTCAGAGCTATC	660						
Db	3600	GGCTACATCAGAGTCTTAGTGATGAAGAAACAGGTATGCGGACAGCTCAGAGCTATC	3659						
Qy	661	TATGACCAAAACCTACGCTGGCGGGGCTGGGTCTATTGTCTTCTCAAGAAATGGTC	720						
Db	3660	TATGACCAAAACCTACGCTGGCGGGGCTGGGTCTATTGTCTTCTCAAGAAATGGTC	3719						
Qy	721	TATTTCTCAGACTCAAGTACGAATGCGAGATATTTAAACAAGATTTGCTCATTTCCG	780						
Db	3720	TATTTCTCAGACTCAAGTACGAATGCGAGATATTTAAACAAGATTTGCTCATTTCCG	3779						
Qy	781	GCAATGCCCTGTGATGCTCAGTCCCTAGACACCTCAGTTCATTTGGTCTCTGGGCT	840						
Db	3780	GCAATGCCCTGTGATGCTCAGTCCCTAGACACCTCAGTTCATTTGGTCTCTGGGCT	3839						
Qy	841	TCTCTCTCTAGCAGCACTTCCTGTCCTTTGACCTTAACCTCTGATGGTTTCCACCTCTG	900						
Db	3840	TCTCTCTCTAGCAGCACTTCCTGTCCTTTGACCTTAACCTCTGATGGTTTCCACCTCTG	3899						
Qy	901	CCAGCAACCCCAACCCAGTGGCTTCAGAGTAAATATCAATGGAATTCAGAGATGAA	960						
Db	3900	CCAGCAACCCCAACCCAGTGGCTTCAGAGTAAATATCAATGGAATTCAGAGATGAA	3959						
Qy	961	CATCTAACCCACTAGAGGAAACAGTTTGGTGATATATGAGACTTTTATGTGGAGTGAAA	1020						
Db	3960	CATCTAACCCACTAGAGGAAACAGTTTGGTGATATATGAGACTTTTATGTGGAGTGAAA	4019						
Qy	1021	TTGGGCAATGCCATTAATGCTTTTCTTTGTTTAAAGAAATGACGTTTACATATA	1080						

Db	4020	TTGGGCAATGCCATTAATGCTTTTCTTTGTTTAAAGAAATGACGTTTACATATA	4079						
Qy	1081	AAATGTAATTTACTTATTGTATTTATGTATATGGAGTTTGAAGGGAATACTGTGCATAAG	1140						
Db	4080	AAATGTAATTTACTTATTGTATTTATGTATATGGAGTTTGAAGGGAATACTGTGCATAAG	4139						
Qy	1141	CCATTATGATAAATAAAGCATGAAAAATATTTGCTGAACTACTTTTGGTGTCTTAAAGTTGT	1200						
Db	4140	CCATTATGATAAATAAAGCATGAAAAATATTTGCTGAACTACTTTTGGTGTCTTAAAGTTGT	4199						
Qy	1201	CACATTCTTGAATTTAGAGTTGCTTACAAATGACACAAATCCCGCTAAATAAATATA	1260						
Db	4200	CACATTCTTGAATTTAGAGTTGCTTACAAATGACACAAATCCCGCTAAATAAATATA	4259						
Qy	1261	AACAAGGCTCAATTTCAAATTTGAAGTAATGTTTGTAGTAAGGAGAGATTAGAAAGCAACAG	1320						
Db	4260	AACAAGGCTCAATTTCAAATTTGAAGTAATGTTTGTAGTAAGGAGAGATTAGAAAGCAACAG	4319						
Qy	1321	GCATAGCAAAATGACATGAAGCTACCGATTAACTAATCGGAACATGTAAACAGTTTACAAA	1380						
Db	4320	GCA TAGCAAAATGACATGAAGCTACCGATTAACTAATCGGAACATGTAAACAGTTTACAAA	4379						
Qy	1381	ATAAACGAATCTCTCTCTTGTCTTCAATGAAGCCCTCATGTGCAGTAGAGATGCAGTT	1440						
Db	4380	ATAAACGAATCTCTCTCTTGTCTTCAATGAAGCCCTCATGTGCAGTAGAGATGCAGTT	4439						
Qy	1441	TCATCAAAAGAAACAAACATCTTGTCAAATGGGTGTGAAGCGGTTCCAGATGTGGATTGGC	1500						
Db	4440	TCATCAAAAGAAACAAACATCTTGTCAAATGGGTGTGAAGCGGTTCCAGATGTGGATTGGC	4499						
Qy	1501	AAAACTCATTTAAGTAAAGTTAGCAGACAAAGTGCAGTGCCTTTAGTGTGCTGCTTGT	1560						
Db	4500	AAAACTCATTTAAGTAAAGTTAGCAGACAAAGTGCAGTGCCTTTAGTGTGCTGCTTGT	4559						
Qy	1561	GCGGTTGTGGGCTGGGGAGGCTCTGCTCAGCTTCCCTCCAGCTTTGCTGCTGCTGAG	1620						
Db	4560	GCGGTTGTGGGCTGGGGAGGCTCTGCTCAGCTTCCCTCCAGCTTTGCTGCTGCTGAG	4619						
Qy	1621	AGGAACACAGACAGACGACAGCGCGGAAAAAGGGCGCATCTAACCGCTATCTAGGCTTTGG	1680						
Db	4620	AGGAACACAGACAGACGACAGCGCGGAAAAAGGGCGCATCTAACCGCTATCTAGGCTTTGG	4679						
Qy	1681	TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTTATTAAGGTTCCAGTTA	1740						
Db	4680	TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTTATTAAGGTTCCAGTTA	4739						
Qy	1741	TAAATATTTTCTTAATATTTTAAAGTCACTATAGAAATGCAACTCCATTTACCACTAAT	1800						
Db	4740	TAAATATTTTCTTAATATTTTAAAGTCACTATAGAAATGCAACTCCATTTACCACTAAT	4799						
Qy	1801	TATTTTAAATATGCTTAGTAAACACATATGTAGTATAATTTCTAGAAAAAACAACATCTAATA	1860						
Db	4800	TATTTTAAATATGCTTAGTAAACACATATGTAGTATAATTTCTAGAAAAAACAACATCTAATA	4859						
Qy	1861	AGTATATAATCTCTGTAATAATAGGCTTGATATATTTAGTTTGTACGATGAGCATG	1920						
Db	4860	AGTATATAATCTCTGTAATAATAGGCTTGATATATTTAGTTTGTACGATGAGCATG	4919						
Qy	1921	CTAGAGCTGTAAACAGAAATACATAGAAATATAGAGGATTTATGATGGAACCTTAATAT	1980						
Db	4920	CTAGAGCTGTAAACAGAAATACATAGAAATATAGAGGATTTATGATGGAACCTTAATAT	4979						
Qy	1981	ATAATGTTGCCAGGATTTTGTCAATATTTTGTACTGTATATCTATCTGTGTATATGG	2040						
Db	4980	ATAATGTTGCCAGGATTTTGTCAATATTTTGTACTGTATATCTATCTGTGTATATGG	5039						
Qy	2041	AATCTCTTTTAAATCAAAACGCTGAAACGAATCAGCATTTAGTCTTGCAGGACACCCAA	2100						
Db	5040	AATCTCTTTTAAATCAAAACGCTGAAACGAATCAGCATTTAGTCTTGCAGGACACCCAA	5099						
Qy	2101	TAACTCAGTCACTGTAAATATGACAAAGTTTGTGTTTTGTTTTGTTTTGTTGTTGGTGG	2160						

QY	1	GAGGACTTGGACGGGTGATGACGGGGTGAATTTTGTAAAGATGATTTTGAATGACAAC	60
Db	3000	GAGGACTTGGACGGGTGATGACGGGGTGAATTTTGTAAAGATGATTTTGAATGACAAC	3059
QY	61	ATCCAGATATTGATGATGTGTCTGCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG	120
Db	3060	ATCCAGATATTGATGATGTGTCTGCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG	3119
QY	121	AACCTCCAGATGTCCTTGGATGCCAAAGGACCAACCAATGATCCCAACTGGGTC	180
Db	3120	AACCTCCAGATGTCCTTGGATGCCAAAGGACCAACCAATGATCCCAACTGGGTC	3179
QY	181	ATTGCGCATCAGGCAAGGAGTGGTTTCAGACGCCAACTCGACCCCGGCTTCTGCTGTA	240
Db	3180	ATTGCGCATCAGGCAAGGAGTGGTTTCAGACGCCAACTCGACCCCGGCTTCTGCTGTA	3239
QY	241	GGTTTTCAGGATTTGGTCTGTGGATCTTCAGTGGGCACTTCAGTAAACACATGACGG	300
Db	3240	GGTTTTCAGGATTTGGTCTGTGGATCTTCAGTGGGCACTTCAGTAAACACATGACGG	3299
QY	301	GACGACGACTATGCTGGCTTGTGTTTGGTTTACAGTCAAGCAGCGCTTCTATGTGGTG	360
Db	3300	GACGACGACTATGCTGGCTTGTGTTTGGTTTACAGTCAAGCAGCGCTTCTATGTGGTG	3359
QY	361	ATGTGGAGCAGGTGACGACGACCTTCTGGGAGGACCAAGCCACGCGGGCCCTATGGCTAC	420
Db	3360	ATGTGGAGCAGGTGACGACGACCTTCTGGGAGGACCAAGCCACGCGGGCCCTATGGCTAC	3419
QY	421	TCCGGCGTGTCCCTCAAGGTGGTGAATCTCAACACGGGACGGGCGAGCACTGAGGAAAC	480
Db	3420	TCCGGCGTGTCCCTCAAGGTGGTGAATCTCAACACGGGACGGGCGAGCACTGAGGAAAC	3479
QY	481	GGCTGTGGCACACGGGAAACACGGCGGGGAGGTGCGAACCTTATGTCACAGACCCACGG	540
Db	3480	GGCTGTGGCACACGGGAAACACGGCGGGGAGGTGCGAACCTTATGTCACAGACCCACGG	3539
QY	541	AACATTGCTGGAAGGACTTACACGGCTTATAGTGGCACCTGACTCAAGGCCCAAGACC	600
Db	3540	AACATTGCTGGAAGGACTTACACGGCTTATAGTGGCACCTGACTCAAGGCCCAAGACC	3599
QY	601	GGCTACATCAGAGTCTTAGTGATGAAGGAAACACAGGTCATGGCAGACTATC	660
Db	3600	GGCTACATCAGAGTCTTAGTGATGAAGGAAACACAGGTCATGGCAGACTATC	3659
QY	661	TATGACCAAACTAGCTGGGGGGGGCTGGGTCTATTTGCTCTCTCAAGAAATGCTC	720
Db	3660	TATGACCAAACTAGCTGGGGGGGGCTGGGTCTATTTGCTCTCTCAAGAAATGCTC	3719
QY	721	TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTGCTGCAATTCGG	780
Db	3720	TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTGCTGCAATTCGG	3779
QY	781	GCAATGCCCTGTGATGCCATGGTCCCTAGACACCTCAGTTCATTTGTTGGTCCCTGGGCT	840
Db	3780	GCAATGCCCTGTGATGCCATGGTCCCTAGACACCTCAGTTCATTTGTTGGTCCCTGGGCT	3839
QY	841	TCTCTCTTAGCAGCAGCTCCCTGCTGCTTGAACCTTCACTGATGTTCTTCACTCTCTG	900
Db	3840	TCTCTCTTAGCAGCAGCTCCCTGCTGCTTGAACCTTCACTGATGTTCTTCACTCTCTG	3899
QY	901	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATCAATGGAATCTCAGAGATGAA	960
Db	3900	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATCAATGGAATCTCAGAGATGAA	3959
QY	961	CATCTAACCCCATAGAGGAAACAGTTTGGTGATATATCAGACTTTTATGAGGAGTGA	1020
Db	3960	CATCTAACCCCATAGAGGAAACAGTTTGGTGATATATCAGACTTTTATGAGGAGTGA	4019
QY	1021	TTGGGATGCCATTAATGCTTTTCTTTTGTGTTTAAAGAAATGAGCTTTTACATATA	1080
Db	4020	TTGGGATGCCATTAATGCTTTTCTTTTGTGTTTAAAGAAATGAGCTTTTACATATA	4079
QY	1081	AAATGTAATTACTTATTTATCTGTATATGAGGTTGAAGGGAATATCTGTGCAATAG	1140
Db	4080	AAATGTAATTACTTATTTATCTGTATATGAGGTTGAAGGGAATATCTGTGCAATAG	4139
QY	1141	CCATTATGATAAATTAAAGCATGAAAAATATTGCTGAACCTACTTTTGGTGTCTTAAAGTGT	1200
Db	4140	CCATTATGATAAATTAAAGCATGAAAAATATTGCTGAACCTACTTTTGGTGTCTTAAAGTGT	4199
QY	1201	CACATTTCTTGAATTAGAGTTGCTTACAAATGACACAAATCCCGCTAAATTAATTAATA	1260
Db	4200	CACATTTCTTGAATTAGAGTTGCTTACAAATGACACAAATCCCGCTAAATTAATTAATA	4259
QY	1261	AACAGGGTCAATTCAAATTTGAAGTAATGTTTTAGTAAAGAGAGATTAGAACACACAG	1320
Db	4260	AACAGGGTCAATTCAAATTTGAAGTAATGTTTTAGTAAAGAGAGATTAGAACACACAG	4319
QY	1321	GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAACAGTTACAAA	1380
Db	4320	GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAACAGTTACAAA	4379
QY	1381	ATAAACGAACTCTCTCTTGTCTTCAATCAATCAAGCCCTCATGTGACGTAGAGATGCAAGTT	1440
Db	4380	ATAAACGAACTCTCTCTTGTCTTCAATCAATCAAGCCCTCATGTGACGTAGAGATGCAAGTT	4439
QY	1441	TCATCAAGAACACAAACATCTTTCGCAATGGGTGTGACGGGTTCACAGATGTGGATTTGGC	1500
Db	4440	TCATCAAGAACACAAACATCTTTCGCAATGGGTGTGACGGGTTCACAGATGTGGATTTGGC	4499
QY	1501	AAAACTCATTTAAGTAAAAAGTTAGCAGAGCAAGTGCCTGCTTTTAGCTGCTGCTTGT	1560
Db	4500	AAAACTCATTTAAGTAAAAAGTTAGCAGAGCAAGTGCCTGCTTTTAGCTGCTGCTTGT	4559
QY	1561	GCCGTTGTGGCGTGGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGCTGAG	1620
Db	4560	GCCGTTGTGGCGTGGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGCTGAG	4619
QY	1621	AGGAACACAGACGACGACGACGCGGAAAGGCGCATCTAACCGCTATCTAGGCTTTGG	1680
Db	4620	AGGAACACAGACGACGACGACGCGGAAAGGCGCATCTAACCGCTATCTAGGCTTTGG	4679
QY	1681	TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA	1740
Db	4680	TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA	4739
QY	1741	TAAATATTTTGTAAATATTTATTAAGTACATATAGAAATGCACTCCATTTACCAGTAACT	1800
Db	4740	TAAATATTTTGTAAATATTTATTAAGTACATATAGAAATGCACTCCATTTACCAGTAACT	4799
QY	1801	TATTTTAAATATGCTAGTAAACATATAGTATAATTTCTAGAAAACAAACATCTAATA	1860
Db	4800	TATTTTAAATATGCTAGTAAACATATAGTATAATTTCTAGAAAACAAACATCTAATA	4859
QY	1861	AGTATATATCTCTGGAATAATAGAGCTTGTATAATTTAGGTTGTCAGATGAAGCATG	1920
Db	4860	AGTATATATCTCTGGAATAATAGAGCTTGTATAATTTAGGTTGTCAGATGAAGCATG	4919
QY	1921	CTAGAAGCTGTAAACAGATAACATAGAGAAATATAGAGGTTTATGATGGAACCTTAATAT	1980
Db	4920	CTAGAAGCTGTAAACAGATAACATAGAGAAATATAGAGGTTTATGATGGAACCTTAATAT	4979
QY	1981	ATAATGTTGCCAGCGATTTTACTTCAATATTTTGTACTGTATCTATCTGCTGTATGCG	2040
Db	4980	ATAATGTTGCCAGCGATTTTACTTCAATATTTTGTACTGTATCTATCTGCTGTATGCG	5039
QY	2041	AAATCTTTTAAATCAAAACGCTGAAACGAAATCAGATTTAGTCTTGCAGGACACACCCAA	2100
Db	5040	AAATCTTTTAAATCAAAACGCTGAAACGAAATCAGATTTAGTCTTGCAGGACACACCCAA	5099
QY	2101	TAAATCAGTCAATGTAATATGACAGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG	2160
Db	5100	TAAATCAGTCAATGTAATATGACAGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG	5159
QY	2161	TTTTTTTGTCTTAAAGTTGATGATCTTTCTGCGAGGAAATAGTCACTCATCCCACTCCACA	2220

```
Db 5160 TTTTTCCTGCTTAAAGTTGCGATGATCTTCTGCGAGAAATAGTCACTCATCCCACTCCACA 5219
Qy 2221 TAAGGGGTTTAGTAAGAGAGAGTCTGCTGCTGCTGATGATGATAGGGGCAAACTTTTTC 2280
Db 5220 TAAGGGGTTTAGTAAGAGAGAGTCTGCTGCTGATGATGATAGGGGCAAACTTTTTC 5279
Qy 2281 CCTTTCTGTTAATAGTCATCACTTCTATGCAAAACAGAGATGCCATACTTTAGT 2340
Db 5280 CCTTTCTGTTAATAGTCATCACTTCTATGCAAAACAGAGATGCCATACTTTAGT 5339
Qy 2341 CTTAATAGTACACATTCGATTTGATATAAATTAATTTGTTGTTTCCCTTTCAGGTTGATCG 2400
Db 5340 CTTAATAGTACACATTCGATTTGATATAAATTAATTTGTTGTTTCCCTTTCAGGTTGATCG 5399
Qy 2401 TTGTTGTTTCTGCTGCACTTTTACTTTTTCGCTGTTGAGCTGTATTTCCCGAGACAAC 2460
Db 5400 TTGTTGTTTCTGCTGCACTTTTACTTTTTCGCTGTTGAGCTGTATTTCCCGAGACAAC 5459
Qy 2461 GAAGCGTTGGGATACCTTCAATTAATAGTACGACTGTCAACAGCGTCAAGGTTTCTGTTT 2520
Db 5460 GAAGCGTTGGGATACCTTCAATTAATAGTACGACTGTCAACAGCGTCAAGGTTTCTGTTT 5519
Qy 2521 CTGTTGTTGGGTCACCGTACAACTGTTGGGATGAGGATGATGTGAATTTAGAA 2580
Db 5520 CTGTTGTTGGGTCACCGTACAACTGTTGGGATGAGGATGATGTGAATTTAGAA 5579
Qy 2581 TGTAACCATATTTTGTAAATTAATTTATGTTTCTTAAACAAATTTATCGTAGTTCGA 2640
Db 5580 TGTAACCATATTTTGTAAATTAATTTATGTTTCTTAAACAAATTTATCGTAGTTCGA 5639
Qy 2641 TGAACCGTCACTGTTTCCCAAGACTGTAAATTTATTTATGTTTCAATGTCACAA 2700
Db 5640 TGAACCGTCACTGTTTCCCAAGACTGTAAATTTATTTATGTTTCAATGTCACAA 5699
Qy 2701 AATTTTCAACCTGAACCTGCACTAGTACAACTTATTTTAAAGATTAACACAGG 2760
Db 5700 AATTTTCAACCTGAACCTGCACTAGTACAACTTATTTTAAAGATTAACACAGG 5759
Qy 2761 AATAAATTTGTAAGGTTTCT 2785
Db 5760 AATAAATTTGTAAGGTTTCT 5784

RESULT 7
HSM807201 3339 bp mRNA linear PRI 28-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKFP686G02190 (from clone DKFP686G02190).
DEFINITION BX641023.1
ACCESSION BX641023.1 GI:34365396
VERSION 1 (bases 1 to 3339)
KEYWORDS Homo sapiens (human)
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3339)
AUTHORS Lauber,J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
CONSRTM Fobo,G., Han,M. and Wiemann,S.
TITLE The German Human cDNA Consortium
JOURNAL Direct Submission
COMMENT Submitted (28-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email a.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFP686G02190) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
location/Qualifiers
1..3339
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="6"
/clone="DKFP686G02190"
/tissue_type="human esophagus tumor"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B_sites SfiIA + SfiIB"
/dev_stage="adult"
1..3339
/gene="DKFP686G02190"
<1..1294
/gene="DKFP686G02190"
/notes="similar to thrombospondin 2, N-terminus truncated"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAE46014.1"
/db_xref="GI:34365397"
/translation="DDNDGVTDKDNKCOLLNFPRQADYDKDEVGDRCDNCPVHNPA
QIDTNNCEGDASVDIDGDDVFNERNDCPYVNTDQRTDGDGVDHDCDNCPLVNP
DQTDVDNLVGPQCDNNEDIDDDGHQNNQDNCPIYSNMQADHRRDGDGACDFFDDN
DQVDDRNCLRVFNPDDQEDLDGDRGDIKDDFDNDIPDIDDDVCPENNAISRTDFR
NFQVPLDLPKGTQTDIPNVIHQKELVQTANSDFGIAVDFEFGVDFSTFYNT
DRDDYAGFYCYQSSRFYVVMKQVITWEDQPTRAYGSGVSLKVVNSTTGTGE
HLRNALWHTGNTPGQVRLTWHDPNIGWKDYTAVERHLTHRPKTVIRVILVHEGQVM
ADSGPIYDQTVAGRLGLFVFSQEMVYFSDLKYECDI"
3305..3310
polyA_signal /gene="DKFP686G02190"
3323
polyA_site /gene="DKFP686G02190"
3323
ORIGIN
Query Match 97.1%; Score 2703; DB 9; Length 3339;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 10; Indels 10; Gaps 5;
Qy 1 GAGGACTTGGACCGTGATGGACGGGTGATATTGTTAAAGATGATTTTGACATGACAAC 60
Db 536 GAGGACTTGGACCGTGATGGACGGGTGATATTGTTAAAGATGATTTTGACATGACAAC 595
Qy 61 ATCCAGATATTGATGATGTGCTCTGAAAAAATGCGATCAGTGAGACAGACTTTCAGG 120
Db 596 ATCCAGATATTGATGATGTGCTCTGAAAAAATGCGATCAGTGAGACAGACTTTCAGG 555
Qy 121 AACTTCAGATGTTCCCTTGGATCCCAAGGACACCCAAATTTGATCCCACTGGGTC 180
Db 656 AACTTCAGATGTTCCCTTGGATCCCAAGGACACCCAAATTTGATCCCACTGGGTC 715
Qy 181 ATTCCCATCAAGGCAAGGAGCTGTTTCAGACAGCAACTCGGACCCCGGCATCGCTGA 240
Db 716 ATTCCCATCAAGGCAAGGAGCTGTTTCAGACAGCAACTCGGACCCCGGCATCGCTGA 775
Qy 241 GGTTTTGAAGTTTGGGCTCTGTGACTTTCAGTGSCACATTTCTAGTAAACACTGACCGG 300
Db 776 GGTTTTGAAGTTTGGGCTCTGTGACTTTCAGTGSCACATTTCTAGTAAACACTGACCGG 835
Qy 301 GACGACGACTATGCTGGCTTCTGTTTGTACAGTCAAGCAGCGCTTCTATGTGTG 360
Db 836 GACGACGACTATGCTGGCTTCTGTTTGTACAGTCAAGCAGCGCTTCTATGTGTG 895
Qy 361 ATGTGGAAGCAGGTGACGACAGCTTCTGCGAGGACACCGGACCGGCTATGCGTAC 420
Db 896 ATGTGGAAGCAGGTGACGACAGCTTCTGCGAGGACACCGGACCGGCTATGCGTAC 955
Qy 421 TCCGCGGTGTTCCCTCAAGGTTGTTGAATCTCCACACGGGAGCGGCGAGCACTGAGGAAC 480
Db 956 TCCGCGGTGTTCCCTCAAGGTTGTTGAATCTCCACACGGGAGCGGCGAGCACTGAGGAAC 1015
Qy 481 CGCTGTGGCACAGGGGAAACACCGCGGGGAGGTGGGAACCTTTATGACACGACCCAGG 540
Db 1016 CGCTGTGGCACAGGGGAAACACCGCGGGGAGGTGGGAACCTTTATGACACGACCCAGG 1075
Qy 541 AACATTGGCTGGAAGGACTACACGGCTTATAGTGGGCACTGACTCAGAGCCCAAGACC 600
```


1076 AACATTGGCTGGAGGACTACAGCGCTATAGTGGCACTGACTCAAGGCCCAAGACT 1135
601 GGCTACATCAGAGTCTTAGTGCATGAGAGAAAACAGGTCTAGGAGACTCAGGACCTATC 660
1136 GGCTACATCAGAGTCTTAGTGCATGAGAGAAAACAGGTCTAGGAGACTCAGGACCTATC 1195
661 TATGACCAAACTACGCTGGGGGGGGCTGGTCTATTTGTTCTCTCAAGAAATGGTC 720
1196 TATGACCAAACTACGCTGGGGGGGGCTGGTCTATTTGTTCTCTCAAGAAATGGTC 1255
721 TATTTCTCAGACCTCAAGTACGAATGACAGAGATATTTAAACAAGATTGTCGATTTCCG 780
1256 TATTTCTCAGACCTCAAGTACGAATGACAGAGATATTTAAACAAGATTGTCGATTTCCG 1315
781 GCAATGCCCTGTGCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
1316 GCAATGCCCTGTGCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1375
841 TCTCTCTAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
1376 TCTCTCTAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1435
901 CCAGCAACCCCAACCCCAAGTCCCTCAGAGGATAAATATCAATGGAATCTCAGAGATGAA 960
1436 CCAGCAACCCCAACCCCAAGTCCCTCAGAGGATAAATATCAATGGAATCTCAGAGATGAA 1495
961 CATCTAACCCCTAGAGAGAACCCAGTGGTGATATAGACCTTATAGTGGAGTGA 1020
1496 CATCTAACCCCTAGAGAGAACCCAGTGGTGATATAGACCTTATAGTGGAGTGA 1555
1021 TTGGGCATGCCATACATGCTTTCTGTTGTTTAAAGAAATGACGTTTACATATA 1080
1556 TTGGGCATGCCATACATGCTTTCTGTTGTTTAAAGAAATGACGTTTACATATA 1615
1081 AAATGTAATTAATTTGATTTATGTTATGTTATGAGTGAAGGAAATCTGTCATAG 1140
1616 AAATGTAATTAATTTGATTTATGTTATGTTATGAGTGAAGGAAATCTGTCATAG 1675
1141 CCAATTAATTAATTAAGCATGAAATAATGCTGAACTCTTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
1676 CCAATTAATTAATTAAGCATGAAATAATGCTGAACTCTTTGTTGTTGTTGTTGTTGTTGTTGTT 1735
1201 CACTATCTTGAATTAAGTCTCTCAAAATGACACACAAATCCCGCTAAATAATA 1260
1736 CACTATCTTGAATTAAGTCTCTCAAAATGACACACAAATCCCGCTAAATAATA 1795
1261 AACAGGGTCAATCAAAATTTGAAGTAAATGTTTATGAGGAGATTTAGAGACACAG 1320
1796 AACAGGGTCAATCAAAATTTGAAGTAAATGTTTATGAGGAGATTTAGAGACACAG 1855
1321 GCATAGCAATGACATAGCTACGATTAATTAATCGGAACATGTAACACAGTTACAAA 1380
1856 GCATAGCAATGACATAGCTACGATTAATTAATCGGAACATGTAACACAGTTACAAA 1915
1381 ATAAAGCAACTCTCTCTGCTCAAAATGAAAGCCCTCATGTGAGTGAAGATGAGT 1440
1916 ATAAAGCAACTCTCTCTGCTCAAAATGAAAGCCCTCATGTGAGTGAAGATGAGT 1975
1441 TCATCAAGAAACAAACATCTTGCATGAAATGGGTGTGACGGGTTCAGATGTTGGC 1500
1976 TCATCAAGAAACAAACATCTTGCATGAAATGGGTGTGACGGGTTCAGATGTTGGC 2035
1501 AAAACCTCAATTAAGTAAAGGTTAGCAGAGCAAGTGGGTGCTTTAGTCTGCTGCT 1560
2036 AAAACCTCAATTAAGTAAAGGTTAGCAGAGCAAGTGGGTGCTTTAGTCTGCTGCT 2095
1561 GCCGTTGTGGCTCGGGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
2096 GCCGTTGTGGCTCGGGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2155
1621 AGGAAACAGAGCAGACGACAGGCCGGAAGGCGCATCTAACGCTATCTAGGCTTTGG 1680

2156 AGGAAACAGAGCAGACGACGCGGGAAGGCGCATCTAACGGTATCTAGGCTTTGG 2215
1681 TAACTCGCGCAAGAGTTCCTTTTACCTGATTTGATGATACATTTTCATTAAGGTTCCAGTTA 1740
2216 TAACTCGCGCAAGAGTTCCTTTTACCTGATTTGATGATACATTTTCATTAAGGTTCCAGTTA 2275
1741 TAAATATTTTGTAAATATTTTAAAGTACATATGATGATGATGATGATGATGATGATGATGATGAT 1800
2276 TAAATATTTTGTAAATATTTTAAAGTACATATGATGATGATGATGATGATGATGATGATGATGAT 2335
1801 TATTTTAAATATGCTAGTAAACATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
2336 TATTTTAAATATGCTAGTAAACATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2395
1861 AGTATATAATCTCTGGAATAATGAGGCTTGAATATTTAGTGTGTCAGGATGAGCATG 1920
2396 AGTATATAATCTCTGGAATAATGAGGCTTGAATATTTAGTGTGTCAGGATGAGCATG 2455
1921 CTAGAAGCTGTAAACAGATACATAGAGATTAATGAGGATTTATGATGGAACCTTAAATA 2515
2456 CTAGAAGCTGTAAACAGATACATAGAGATTAATGAGGATTTATGATGGAACCTTAAATA 2575
1980 TATAATGTTGCGCAGGATTTTGTGTTCAATATTTGTTACTGTTATCTATCTGCTGTATATG 2039
2516 TATAATGTTGCGCAGGATTTTGTGTTCAATATTTGTTACTGTTATCTATCTGCTGTATATG 2575
2040 GAAATCTTTTAAATCAACCGCTGAAACGAAATGAGCAATTTAGTGTGTCAGGACACACCA 2099
2576 GAAATCTTTTAAATCAACCGCTGAAACGAAATGAGCAATTTAGTGTGTCAGGACACACCA 2634
2100 ATAATCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2159
2635 ATAATCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2694
2160 ----GTTTGTGTTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2215
2695 GTTTGTGTTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2754
2216 CCACATAAGGGGTTTGTAGTAAAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2814
2755 CCACATAAGGGGTTTGTAGTAAAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2874
2276 TTTTCCCTCTGTTTAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2335
2815 TTTTCCCTCTGTTTAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2874
2336 TTAGTCTTAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2395
2875 TTAGTCTTAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2934
2396 GATCGTTG---TGTTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2452
2935 GATCGTT 2994
2453 GAGA--CAACGAAGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2511
2995 GAGACCAAGAGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3054
2512 TTTCTGTTCTGTT 2571
3055 TTTCTGTTCTGTT 3114
2572 TATTTAGAT 2631
3115 TATTTAGAT 3174
2632 ATAGTGTGAT 2691
3175 ATAGTGTGAT 3234
2692 CATGTTCAAAATTTCAACCACTGAAACCCCTGACCTTAGCTAGAACCTTCAATTTTAAAGATT 2751
3235 CATGTTCAAAATTTCAACCACTGAAACCCCTGACCTTAGCTAGAACCTTCAATTTTAAAGATT 3294


```

QY 2752 AACACAGGAAATAAATTTGTAAGAAAGGT 2780
Db 3295 AACACAGGAAATAAATTTGTAAGAAAGGT 3323

RESULT 8
BV177227
LOCUS
DEFINITION sqm93505 Human DNA (Sequenom) Homo sapiens STS 10-JUN-2004
tagged site.
ACCESSION BV177227
VERSION BV177227.1 GI:48013259
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5505)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
Fax: 18582029020
Email: abraun@sequenom.com
Primer A: No primer sequence submitted
Primer B: No primer sequence submitted
STS size: 5505.

FEATURES
source
1..5505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human DNA (Sequenom)"
<1..>5505

STS
ORIGIN
Query Match 85.2%; Score 2372.8; DB 11; Length 5505;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2469; Conservative 1; Mismatches 29; Indels 9; Gaps 8;

QY 1 GAGGACTTGGACGGTGATGGACGGGGTGATA-TTTGTAAGATGATTTTGACAATGACAA 59
Db 3003 GAGGACTTGGACGGTGATGGACGGGGTGATA-TTTGTAAGATGATTTTGACAATGACAA 3062

QY 60 CATCCAGATATTGATGATGTGTCTCGAAACAAATGCCATCATGAGACAGACTTCA- 118
Db 3063 CATCCAGATATTGATGATGTGTCTCGAAACAAATGCCATCATGAGACAGACTTCCAG 3122

QY 119 GGAATCTCCAGATGTCCTTGGATCCCAAGGAGACCAACCAATTCATCCCACTGGG 178
Db 3123 GGAATCTCCAGATGTCCTTGGATCCCAAGGAGACCAACCAATTCATCCCACTGGG 3182

QY 179 TCATTGCGCATCAGGCAAGGAGCTGTTCAGACAGCAATTCGGACCCCGGATCGCTG 238
Db 3183 TCATTGCGCATCAGGCAAGGAGCTGTTCAGACAGCAATTCGGACCCCGGATCGCTG 3242

QY 239 TAGGTTTTGACGAGTTTGGGCTGTGGACTTCAGTGGCAGACTTCTACGTAACACTGACC 298
Db 3243 TAGGTTTTGACGAGTTTGGGCTGTGGACTTCAGTGGCAGACTTCTACGTAACACTGACC 3302

QY 299 GGGACGACGACTATGCTGGCTTCGTTCTTTGGTTACAGTCAAGCAGCGCTTCATATGG 358
Db 3303 GGGACGACGACTATGCTGGCTTCGTTCTTTGGTTACAGTCAAGCAGCGCTTCATATGG 3362

QY 359 TGATGTGAAGAGGTGACGAGACCTTACTCGGAGGACCAAGCCACGCGGGCCTATGGCT 418

```

```

Db 3363 TGATGTGAAGAGGTGACGAGACCTTACTGGAGAGACAGCCACGCGGC--TATGGT 3420
QY 419 ACTCCGGCGTGTCCCTCAAGGTGGTGAATCCACCAACGGGGACGGCGAGCAGCTGAGGA 478
Db 3421 ACTCCGGCGTGTCCCTCAAGGTGGTGAATCCACCAACGGGGACGGCGAGCAGCTGAGGA 3480
QY 479 AGCGCTGTGGCACACCGGGGAACACCGCGGGGAGGTGCGAACCCTTATGGCAGACCCCA 538
Db 3481 AGCGCTGTGGCACACCGGGGAACACCGCGGGGAGGTGCGAACCCTTATGGCAGACCCCA 3540
QY 539 GGAACATTCGCTGGAAGGACTACACGGCTATAGGTGGCACCCTGACTCAGACGCCCAAGA 598
Db 3541 GGAACATTCGCTGGAAGGACTACACGGCTATAGGTGGCACCCTGACTCAGACGCCCAAGA 3600
QY 599 CGCGCTATCATGAGACTTATGATGATGAAGAAACAGGTCTATGGCAGACTCAGGACCTTA 658
Db 3601 CGCGCTATCATGAGACTTATGATGATGAAGAAACAGGTCTATGGCAGACTCAGGACCTTA 3660
QY 659 TCTATGACCAACCTACGCTGGCGGGCGCTGGTCTATTGCTCTCTCAAGAAATGG 718
Db 3661 TCTATGACCAACCTACGCTGGCGGGCGCTGGTCTATTGCTCTCTCAAGAAATGG 3720
QY 719 TCTATTTCTCAGACCTCAAGTACAGATATTTAAACAAGATTTGCTGCAATTC 778
Db 3721 TCTATTTCTCAGACCTCAAGTACAGATATTTAAACAAGATTTGCTGCAATTC 3780
QY 779 CGGCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTCATTTGTCCTTCGCG 838
Db 3781 CGGCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTCATTTGTCCTTCGCG 3840
QY 839 CTTCTCTCTAGCAGACCTCTCTGCTGACCTTAACCTGATGCTGATGCTTCACTCTCC 898
Db 3841 CTTCTCTCTAGCAGACCTCTCTGCTGACCTTAACCTGATGCTGATGCTTCACTCTCC 3900
QY 899 TGCCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATATCAATGAACCTCAGAGATG 958
Db 3901 TGCCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATATCAATGAACCTCAGAGATG 3960
QY 959 AACATCTAACCCACTAGAGGAAACCCAGTTTGGTGATATATGAGACTTTATGTGGAGTGA 1018
Db 3961 AACATCTAACCCACTAGAGGAAACCCAGTTTGGTGATATATGAGACTTTATGTGGAGTGA 4020
QY 1019 AATTGGGCGATGCCATTACATTTGCTTTTCTGTTGTTTAAAGAAATGACCTTTACATA 1078
Db 4021 AATTGGGCGATGCCATTACATTTGCTTTTCTGTTGTTTAAAGAAATGACCTTTACATA 4080
QY 1079 TAAATGTAAATTACTTATTGTTATTTATGTTATATGAGTTTGAAGGAAATACTGTGCATA 1138
Db 4081 TAAATGTAAATTACTTATTGTTATTTATGTTATATGAGTTTGAAGGAAATACTGTGCATA 4140
QY 1139 AGCCATTATGATAAATAAGCATGAAAATAATTCGTAACCTACTTTTGGTGTCTTAAAGTT 1198
Db 4141 AGCCATTATGATAAATAAGCATGAAAATAATTCGTAACCTACTTTTGGTGTCTTAAAGTT 4200
QY 1199 GTCACTATTCTTGAATTAGAGTTGCTCTAATGACACACAATAATCCGCTTAATAAATTA 1258
Db 4201 GTCACTATTCTTGAATTAGAGTTGCTCTAATGACACACAATAATCCGCTTAATAAATTA 4260
QY 1259 TAAACAAGGGTCAATTCAAATTTGAAGTAATGTTTATGAGGAGAGATTAGAGACAAC 1318
Db 4261 TAAACAAGGGTCAATTCAAATTTGAAGTAATGTTTATGAGGAGAGATTAGAGACAAC 4320
QY 1319 AGGCATAGCAATAGCATAAGCTACCGATTAACTCGGAACATGTAAACAGTTACAA 1378
Db 4321 AGGCATAGCAATAGCATAAGCTACCGATTAACTCGGAACATGTAAACAGTTACAA 4380
QY 1379 AAATAAAGCACTCTCTCTGCTTCAATGAAGCCCTCATGTGAGTAGAGATGAG 1438
Db 4381 AAATAAAGCACTCTCTCTGCTTCAATGAAGCCCTCATGTGAGTAGAGATGAG 4440
QY 1439 TTTTCATCAAGAAACAAACATCTTGTGCAATGGGTGTCACCGGTTCCAGATGTGATTTG 1498

```

Db	4441	TTTCAATCAAGAAACAAACATCTTTGCAAAATGGGTGTGANGCGGTTCAGATGTGGATTG	4500
Qy	1499	GCAAAACCTCATTTAAAGTAAAGGTTAGCAGAGCAAAAGTGGGTCTTTAGCTCTCTGCTT	1558
Db	4501	GCAAAACCTCATTTAAAGTAAAGGTTAGCAGAGCAAAAGTGGGTCTTTAGCTCTCTGCTT	4560
Qy	1559	GTGCGGTGTGGCTGGGAGGCTCTGCTGAGCTTCCTTCCCGAGCTTTGCTGCTG	1618
Db	4561	GTGCGGTGTGGCTGGGAGGCTCTGCTGAGCTTCCTTCCCGAGCTTTGCTGCTG	4620
Qy	1619	AGAGAAACAGAGCAGACGCAAGCGCGGAAAGCGGATCTAACGGGTATCTAGGCTTT	1678
Db	4621	AGAGAAACAGAGCAGACGCAAGCGCGGAAAGCGGATCTAACGGGTATCTAGGCTTT	4680
Qy	1679	GGTAACCTGCGACAGTTGCTTTTACCTGATTTTCATATGATATATTTTCAATAGGTTCCAGT	1738
Db	4681	GGTAACCTGCGACAGTTGCTTTTACCTGATTTTCATATGATATATTTTCAATAGGTTCCAGT	4740
Qy	1739	TATAAATATTTTGTAAATATTTTAAAGTACTATAGAAATCAATCCATTTACAGTAA	1798
Db	4741	TATAAATATTTTGTAAATATTTTAAAGTACTATAGAAATCAATCCATTTACAGTAA	4800
Qy	1799	CTTATTTTAAATATGCTTAGTAAACATATGATATATTTTCTAGAAACAAACATCTAA	1858
Db	4801	CTTATTTTAAATATGCTTAGTAAACATATGATATATTTTCTAGAAACAAACATCTAA	4860
Qy	1859	TAGTATATATATCTGCGAAATATGAGCTTGTGATAATATTAGGTTGTCAGATGAAGCA	1918
Db	4861	TAGTATATATATCTGCGAAATATGAGCTTGTGATAATATTAGGTTGTCAGATGAAGCA	4919
Qy	1919	TGCTAGAAAGCTGTAAACAGAAATACATAGAGAAATATGAGAGTTTATGATGAA-CCCTTAA	1977
Db	4920	TGCT-AGAGCTGT-ACCGAATACATAGAGAAATATGAGAGTTTATGATGAA-CCCTTAA	4977
Qy	1978	TATATAATGTTGCGAGCAATTTAGTCAATATTTTACTGTTATCTATCTGCTGATA	2037
Db	4978	TATATAATGTTGCGAGCAATTTAGTCAATATTTTACTGTTATCTATCTGCTGATA	5037
Qy	2038	TGGAATCTTTTAAATCAAAAGCTGAAACAGAAATGAGCTTTAGTCTTCCAGGACACACC	2097
Db	5038	TGGAATCTTTTAAATCAAAAGCTGAAACAGAAATGAGCTTTAGTCTTCCAGGACACACC	5097
Qy	2098	CAATAATCAGTCATGTGTAATATGCAAAAGTTTGTGTTTTTGTGTTTTTGTGTTGT	2157
Db	5098	CAATAATCAGTCATGTGTAATATGCAAAAGTTTGTGTTTTTGTGTTTTTGTGTTGT	5157
Qy	2158	TGGTTTTTTTGTCTTAAAGTTCATGATCTTTCTGAGGAAATAGTCACTATCCCACTCC	2217
Db	5158	TGGTTTTTTTGTCTTAAAGTTCATGATCTTTCTGAGGAAATAGTCACTATCCCACTCC	5217
Qy	2218	ACATAAGGGTTTGTAGAGAGAGTCTGTCTGCTGATGATGATAGGGGGCAAAATCTTT	2277
Db	5218	ACATAAGGGTTTGTAGAGAGAGTCTGTCTGCTGATGATGATAGGGGGCAAAATCTTT	5277
Qy	2278	TTCCCTTTCTGTTAATAGTCAATCTTATGACCAACAGGAACGATCCATTAATCTTT	2337
Db	5278	TTCCCTTTCTGTTAATAGTCAATCTTATGACCAACAGGAACGATCCATTAATCTTT	5337
Qy	2338	AGTCTTAATGTACACATTTGATATAAATTAATTTGTTTGTGTTTTGTTGAGTTGA	2397
Db	5338	AGTCTTAATGTACACATTTGATATAAATTAATTTGTTTGTGTTTTGTTGAGTTGA	5397
Qy	2398	TGCTGTGTTGTTTGTCTGACCTTTTATCTTTTGTGCGTGTGAGCTGTATTTCCCGAGA-	2456
Db	5398	TGCTGTGTTGTTTGTCTGACCTTTTATCTTTTGTGCGTGTGAGCTGTATTTCCCGAGAC	5457
Qy	2457	CAACGAAGGCTTGGGATATTTCAATTAATGTAGCGACTGTCAACAGCG	2504
Db	5458	CAACGAAGGCTTGGGATATTTCAATTAATGTAGCGACTGTCAACAGCG	5505
RESULT 9			
BV177983			

LOCUS	BV177983	sgnm97822	Human DNA (Sequenom)	5505 bp	DNA	linear	STS 10-JUN-2004					
DEFINITION	tagged site.											
ACCESSION	BV177983											
VERSION	BV177983.1	GI:48014218										
KEYWORDS	STS.											
SOURCE	Homo sapiens (human)											
ORGANISM	Homo sapiens											
REFERENCE	1 (bases 1 to 5505)											
AUTHORS	Nelson, R.M., Marnellos, G., Kammerer, S., Hoyal, C.R., Shi, M.M., Cantor, C.R., and Braun, A.											
TITLE	Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions											
JOURNAL	Genome Res. (2004) In press											
COMMENT	Contact: Andreas Braun Pharmaceuticals division Sequenom, Inc. 3595 John Hopkins Court, San Diego, CA 92121, USA Tel: 18582029018 Fax: 18582029020 Email: abraun@sequenom.com Primer A: No primer sequence submitted Primer B: No primer sequence submitted STS size: 5505.											
FEATURES	Location/Qualifiers											
source	1..5505											
	/organism="Homo sapiens"											
	/mol_type="genomic DNA"											
	/db_xref="taxon:9606"											
	/clone_lib="Human DNA (Sequenom)"											
	<1..>5505											
STS												
ORIGIN												
Query Match	85.2%; Score 2372.8; DB 11; Length 5505;											
Best Local Similarity	98.4%; Pred. No. 0;											
Matches 2469; Conservative	1; Mismatches 29; Indels 9; Gaps 8;											
Qy	1	GAGGACTTGGACGGTGTGATGGACGGGGTGATA-TTTGTAAGATGATTTTGACAAATGACAA	59									
Db	3003	GAGGACTTGGACGGTGTGATGGACGGGGTGATA-TTTGTAAGATGATTTTGACAAATGACAA	3062									
Qy	60	CATCCAGATATTTGATGATGTGTCTGAAAACAAATGCCATCAGTGACAGACATTCATCA-118										
Db	3063	CATCCAGATATTTGATGATGTGTCTGAAAACAAATGCCATCAGTGACAGACATTCATCA-3122										
Qy	119	GGAACTTCCAGATGTGTCTGAAAACAAATGCCATCAGTGACAGACATTCATCACTGGG178										
Db	3123	GGAACTTCCAGATGTGTCTGAAAACAAATGCCATCAGTGACAGACATTCATCACTGGG3182										
Qy	179	TCATTTCGCGCATCAAGCAAGGAGCTGGTTTCAGACAGCAACTCGGACCCCGGATCGCTG238										
Db	3183	TCATTTCGCGCATCAAGCAAGGAGCTGGTTTCAGACAGCAACTCGGACCCCGGATCGCTG3242										
Qy	239	TAGGTTTTCAGAGTTTGGTCTGTGGACTTCAGTGGGACATTTCTACGTAAACACTGACC298										
Db	3243	TAGGTTTTCAGAGTTTGGTCTGTGGACTTCAGTGGGACATTTCTACGTAAACACTGACC3302										
Qy	299	GGGACGACGACTATGCTGCTTTCGTTTACAGTCAAGCAGCGCGCTTCTATGTGG358										
Db	3303	GGGACGACGACTATGCTGCTTTCGTTTACAGTCAAGCAGCGCGCTTCTATGTGG3362										
Qy	359	TGATGTGGAAGCAGGTTGACGACGACTACTCGGAGGACCAAGCCACGCGGGGCTATGGCT418										
Db	3363	TGATGTGGAAGCAGGTTGACGACGACTACTCGGAGGACCAAGCCACGCGGGGCTATGGCT3420										
Qy	419	ACTCCGGCTGTCTCTCAAGGTGGTGAATCCACACGGGCGAGCGGCGAGACCTCAGGA478										
Db	3421	ACTCCGGCTGTCTCTCAAGGTGGTGAATCCACACGGGCGAGCGGCGAGACCTCAGGA3480										
Qy	479	ACGCGCTGTGGCACACGGGGAACACACCGCGGGGAGGTGCGAACTTTATGGCAGCACCCCA538										

|||||
3481 AC CGCGTGTGGCACACGGGGAAACCGCGGGGAGGTGGAACTTTATGGCAGCCCCA 3540
QY
539 GGAACATTGCGTGGAGGACTACACGGCTATAGTGGGACCTGACTCAGGCCCAAGA 598
Db
3541 GGAACATTGCGTGGAGGACTACACGGCTATAGTGGGACCTGACTCAGGCCCAAGA 3600
QY
599 CGGCTACATCAGAGTCTTAGTGCATGAAGGAAACAGAGTGCATGGCAGACTCAGGACCTA 658
Db
3601 CTGGCTACATCAGAGTCTTAGTGCATGAAGGAAACAGAGTGCATGGCAGACTCAGGACCTA 3660
QY
659 TCTATGACCAACCTACGCTGGCGGGCGGTGGGTCTATTGTCTCTCAAGAAATGG 718
Db
3661 TCTATGACCAACCTACGCTGGCGGGCGGTGGGTCTATTGTCTCTCAAGAAATGG 3720
QY
719 TCTATTCTCAGACCTCAAGTACGATGACAGATATTTAAACAGATTTGCTGCATTTTC 778
Db
3721 TCTATTCTCAGACCTCAAGTACGATGACAGATATTTAAACAGATTTGCTGCATTTTC 3780
QY
779 CGGCAATGCCCTGTGCATGCGCATGGTCCCTAGACACCTCAGTTCAATTTGGTGCCTTGGCG 838
Db
3781 CGGCAATGCCCTGTGCATGCGCATGGTCCCTAGACACCTCAGTTCAATTTGGTGCCTTGGCG 3840
QY
839 CTTCTCTCTAGCAGACCTCTCTGCTCCCTTGACCTTAACTCTGATGGTCTTCACTCC 898
Db
3841 CTTCTCTCTAGCAGACCTCTCTGCTCCCTTGACCTTAACTCTGATGGTCTTCACTCC 3900
QY
899 TGCCAGCAACCCCAACCAAGTCCCTCAGAGGATAAATCAATGGAACCTCAGAGATG 958
Db
3901 TGCCAGCAACCCCAACCAAGTCCCTCAGAGGATAAATCAATGGAACNOCAGAGATG 3960
QY
959 AACATCTAAACCCACTAGAGGAAACCAAGTTTGGTGATATATAGACTTTATGTGGAGTGA 1018
Db
3961 AACATCTAAACCCACTAGAGGAAACCAAGTTTGGTGATATATAGACTTTATGTGGAGTGA 4020
QY
1019 AATTGGGCATGCCATTA CATTGCTTTTCTGTTTGTGTTTAAAAAGAACGCTTTACATA 1078
Db
4021 AATTGGGCATGCCATTA CATTGCTTTTCTGTTTGTGTTTAAAAAGAACGCTTTACATA 4080
QY
1079 TAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1138
Db
4081 TAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4140
QY
1139 AGCCATTATGATAAATTAAGCATGAAAAATATGCTGAACTACTTTTGGTGTCTTAAAGTT 1198
Db
4141 AGCCATTATGATAAATTAAGCATGAAAAATATGCTGAACTACTTTTGGTGTCTTAAAGTT 4200
QY
1199 GTCACTATCTTGAATTAGAGTTGCTCTCAATGACACACAAATCCCGCTAAATAATTA 1258
Db
4201 GTCACTATCTTGAATTAGAGTTGCTCTCAATGACACACAAATCCCGCTAAATAATTA 4260
QY
1259 TAAACAAGGGTCAATTCAAATTTGAAGTAATGTTTATAGTAAAGGAGAGATTAGAAGCAAC 1318
Db
4261 TAAACAAGGGTCAATTCAAATTTGAAGTAATGTTTATAGTAAAGGAGAGATTAGAAGCAAC 4320
QY
1319 AGGCATAGCAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAACAGTTACAA 1378
Db
4321 AGGCATAGCAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAACAGTTACAA 4380
QY
1379 AAATAAAGCACTCTCTCTTGTCTTCTCAATGAAAGCCCTCATGTGCAGTAGAGATGCAG 1438
Db
4381 AAATAAAGCACTCTCTCTTGTCTTCTCAATGAAAGCCCTCATGTGCAGTAGAGATGCAG 4440
QY
1439 TTTTCATCAAGAAACAAACATCTTCAATGAGTGGTGAGCGGTTTCCAGATGTGATTTG 1498
Db
4441 TTTTCATCAAGAAACAAACATCTTCAATGAGTGGTGAGCGGTTTCCAGATGTGATTTG 4500
QY
1499 GCAAAACCTCAITTAAGTAAAGGTTAGCAGAGCAAAAGTGGCGTCTTTAGCTGTGCTT 1558
Db
4501 GCAAAACCTCAITTAAGTAAAGGTTAGCAGAGCAAAAGTGGCGTCTTTAGCTGTGCTT 4560
QY
1559 GTGCGGTTGTGGCGTGGGGAGGCTCTGCTGAGCTTCTTCCCGAGCTTTGTGCGCTG 1618
|||||

Db
4561 GTGCGCTGTGGNGTCGGGAGGCTCTCGCTGAGCTTCTTCCCGAGCTTTGTCTGCTG 4620
QY
1619 AGAGGAACAGAGCAGAGCAGCAGCGCGGAAAGGCGCATCTAAACGGTATCTAGGCTTT 1678
Db
4621 AGAGGAACAGAGCAGAGCAGCAGCGCGGAAAGGCGCATCTAAACGGTATCTAGGCTTT 4680
QY
1679 GGTAACTGCGGACAACTGCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGT 1738
Db
4681 GGTAACTGCGGACAACTGCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGT 4740
QY
1739 TATAAATATTTTGTAAATATTTTAAAGTGAATATAGAAATGCAATCTCCATTTTACCAGTAA 1798
Db
4741 TATAAATATTTTGTAAATATTTTAAAGTGAATATAGAAATGCAATCTCCATTTTACCAGTAA 4800
QY
1799 CTTATTTTAAATATGCTTAGTAACACATATGTAGTATAATTTCTAGAAACAAACATCTCAA 1858
Db
4801 CTTATTTTAAATATGCTTAGTAACACATATGTAGTATAATTTCTAGAAACAAACATCTCAA 4860
QY
1859 TAAAGTATATAAATCTGCTGAAAAATATGAGGCTTTGATAATATAGGTTGTCCAGATGAAGCA 1918
Db
4861 TAAAGTATATAAATCTGCTGAAAAATATGAGGCTTTGATAATATAGGTTGTCCAGATG-AGCA 4919
QY
1919 TCCTAGAAAGCTGTAAACAGAAATACATAGAGAAATAATGAGGAGTTTATGATGGAA-CCTTAA 1977
Db
4920 TGCT-GAAGCTGT-ACCGAATACATAGAGAAATAATGAGAGTGTATGATGGAACCTTAA 4977
QY
1978 TATATAATGTTGCCAGCGAATTTAGTTCAATATTTGTTACTGTTATCTATCTGTGTATA 2037
Db
4978 TATATAATGTTGCCAGCGAATTTAGTTCAATATTTGTTACTGTTATCTATCTGTGTATA 5037
QY
2038 TCGAATTTCTTTTAAATTTCAAACTGAAACGAAATCAGCAATTTAGTCTTCCAGGACACCC 2097
Db
5038 TCGAATTTCTTTTAAATTTCAAACTGAAACGAAATCAGCAATTTAGTCTTCCAGGACACCC 5097
QY
2098 CAATAATCAGTCATGTGTAATAATGCACAAGTTTGTGTTTTTGTGTTTTTGTGTTGT 2157
Db
5098 CAATAATCAGTCATGTGTAATAATGCACAAGTTTGTGTTTTTGTGTTTTTGTGTTGT 5157
QY
2158 TGGTTTTTGTGTTTAAAGTTGATGATCTTCTGCAAGAAATAGTCACTCATCCACTCC 2217
Db
5158 TGGTTTTTGTGTTTAAAGTTGATGATCTTCTGCAAGAAATAGTCACTCATCCACTCC 5217
QY
2218 ACATAAGGGTTTTAGTAAAGAGAGTCTGTCTGTGATGATGATAGGGGCAAAATCTTT 2277
Db
5218 ACATAAGGGTTTTAGTAAAGAGAGAGTCTGTCTGTGATGATGATAGGGGCAAAATCTTT 5277
QY
2278 TTCCCTTTCTGTTTAAATAGTCACTCATCTTCTATGCCAAACAGGAACGATCACTAACTTT 2337
Db
5278 TTCCCTTTCTGTTTAAATAGTCACTCATCTTCTATGCCAAACAGGAACGATCACTAACTTT 5337
QY
2338 AGTCTTAATGACACATTTGCAATTTTGAATAAATTTTGTGTTTCTTTGAGTTGA 2397
Db
5338 AGTCTTAATGACACATTTGCAATTTTGAATAAATTTTGTGTTTCTTTGAGTTGA 5397
QY
2398 TCGTTGTGTTGTTTGTGTCACCTTTTACTTTTTTGTGCGTGTGAGCTGTATTTCCCGAG- 2456
Db
5398 TCGTTGTGTTGTTTGTGTCACCTTTTACTTTTTTGTGCGTGTGAGCTGTATTTCCCGAG- 5457
QY
2457 CAACGAAGCGTTGGGATACCTTCAATTAATGTAGCGACTGTCAACAGCG 2504
Db
5458 CAACGAAGCGTTGGGATACCTTCAATTAATGTAGCGACTGTCAACAGCG 5505
RESULT 10
BVI78635
LOCUS
DEFINITION
sqm101716 Human DNA (Sequenc) Homo sapiens STS genomic, sequence
tagged site.
ACCESSION
BVI78635
VERSION
BVI78635.1 GI:48015079
KEYWORDS
STS,
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

Db 4681 GGTAACGCGGCAAGTGTGCTTTTACCTGATTTGATGATACATTTTCATTAAGGTTCCAGT 4740

QY 1739 TATAAATATTTTCTTAATAATTTTAAAGTGACATATAGAAATGCAACTCCATTTTACCAGTAA 1798

Db 4741 TATAAATATTTTGTATTAATTTTAAAGTGACATATAGAAATGCAACTCCATTTTACCAGTAA 4800

QY 1799 CTTATTTTAAATATGCTAGTAAACATATGTAGTATAATTTCTAGAAACAAACATCTAA 1858

Db 4801 CTTATTTTAAATATGCTAGTAAACATATGTAGTATAATTTCTAGAAACAAACATCTAA 4860

QY 1859 TAAGTATATAATCCCTGTGAAATATAGGCTTGTAGTATAATTTCTAGAAACAAACATCTAA 1918

Db 4861 TAAGTATATAATCCCTGTGAAATATAGGCTTGTAGTATAATTTCTAGAAACAAACATCTAA 4919

QY 1919 TGCTAGAACTGTAAACAGAAATACATAGAAATATAGGAGTATTTATGATGAA-CCTTAA 1977

Db 4920 TGCT-GAAGCTGT-ACCGAATACATAGAAATATAGGAGTATTTATGATGAAACCTTTAA 4977

QY 1978 TATATAATGTTGCCAGGATTTTGTGTCATATTTTGTACTGTATCTATCTGCTATA 2037

Db 4978 TATATAATGTTGCCAGGATTTTGTGTCATATTTTGTACTGTATCTATCTGCTATA 5037

QY 2038 TGGAAATCTTTTAAATCAACGCTGAAACGATCAGCAATTTAGTCTTCCGAGGACACC 2097

Db 5038 TGGAAATCTTTTAAATCAACGCTGAAACGATCAGCAATTTAGTCTTCCGAGGACACC 5097

QY 2098 CAATAATCAGTCATGTATATATGCAAAAGTTTGTGTTTTTGTGTTTTTGTGTTGT 2157

Db 5098 CAATAATCAGTCATGTATATATGCAAAAGTTTGTGTTTTTGTGTTTTTGTGTTGT 5157

QY 2158 TGGTTTTTGTGTTTAAAGTTGATGATCTTTCTGAGGAAATAGTCACTCATCCACTCC 2217

Db 5158 TGGTTTTTGTGTTTAAAGTTGATGATCTTTCTGAGGAAATAGTCACTCATCCACTCC 5217

QY 2218 ACATAAGGGTTTGTAGGAGAGAGTCTGTCTGTGATGATGATGAGGAGGCAATCTTT 2277

Db 5218 ACATAAGGGTTTGTAGGAGAGAGTCTGTCTGTGATGATGATGAGGAGGCAATCTTT 5277

QY 2278 TTCCCTCTTCTGTATATAGTCACTCATCTTCTATGCAACAGGAGGATCCATAACTTT 2337

Db 5278 TTCCCTCTTCTGTATATAGTCACTCATCTTCTATGCAACAGGAGGATCCATAACTTT 5337

QY 2338 AGCTTAATGTACATGTCATTTTGTATAAAATTAATTTGTTTCTTTCAGGTTGA 2397

Db 5338 AGCTTAATGTACATGTCATTTTGTATAAAATTAATTTGTTTCTTTCAGGTTGA 5397

QY 2398 TCGTTGTGTTTGTGCTGCACTTTTACTTTTTCGCGTGTGAGCTGTATTCGAGAG- 2456

Db 5398 TCGTTGTGTTTGTGCTGCACTTTTACTTTTTCGCGTGTGAGCTGTATTCGAGAG 5457

QY 2457 CAACGAAGCGTTGGGATCTTCAATTAATGTAGCGACTGTCAACAGCG 2504

Db 5458 CAACGAAGCGTTGGGATCTTCAATTAATGTAGCGACTGTCAACAGCG 5505

RESULT 11

LOCUS BV179851

DEFINITION sqm106999 Human DNA (Sequenc) Homo sapiens STS genomic, sequence tagged site.

ACCESSION BV179851

VERSION BV179851.1 GI:48016717

KEYWORDS STS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5505)

AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.

TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions

JOURNAL Genome Res. (2004) In press

COMMENT

Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
Fax: 18582029020
Email: abraun@sequenom.com

Primer A: No primer sequence submitted
Primer B: No primer sequence submitted
STS size: 5505.

FEATURES

Location/Qualifiers
1..5505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human DNA (Sequenc)"
<1..5505

STS
ORIGIN

Query Match 85.2%; Score 2372.8; DB 11; Length 5505;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2469; Conservative 1; Mismatches 29; Indels 9; Gaps 8;

QY 1 GAGGACTTGGACGGTGATGGACGGGTGATA-TTTGTAAGATGATTTTGACAATGACAA 59

Db 3003 GAGGACTTGGACGGTGATGGACGGGTGATA-TTTGTAAGATGATTTTGACAATGACAA 3062

QY 60 CATCCAGATATTGATGATGTGTCTCTGAAACAAATGCCATCAGTGAGACAGACTTCA- 118

Db 3063 CATCCAGATATTGATGATGTGTCTCTGAAACAAATGCCATCAGTGAGACAGACTTCA 3122

QY 119 GGAATCTCCAGATGTCTCCCTTTGGATCCCAAGGAGACACCCAAATTTGATCCCACTGGG 178

Db 3123 GGAATCTCCAGATGTCTCCCTTTGGATCCCAAGGAGACACCCAAATTTGATCCCACTGGG 3182

QY 179 TCATTTCGCCATCAAGCGAAGGAGCTGGTTTCAGACAGCAACTCGGACCCCGGCGATCGCTG 238

Db 3183 TCATTTCGCCATCAAGCGAAGGAGCTGGTTTCAGACAGCAACTCGGACCCCGGCGATCGCTG 3242

QY 239 TAGGTTTTGACGAGTTTGGGTCTGTGGACTTCAGTGGCACATTTACGTAACACACTGACC 298

Db 3243 TAGGTTTTGACGAGTTTGGGTCTGTGGACTTCAGTGGCACATTTACGTAACACACTGACC 3302

QY 299 GGGACGACGACTATGCTGCTGCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGG 358

Db 3303 GGGACGACGACTATGCTGCTGCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGG 3362

QY 359 TGATGTGGAAGCAGGTGACGACGACTTCTGGGAGGACGAGCCACGCGGGCTTATGGCT 418

Db 3363 TGATGTGGAAGCAGGTGACGACGACTTCTGGGAGGACGAGCCACGCGGGCTTATGGCT 3420

QY 419 ACTCCGGCTGTCTCAAGGTGGTGAATCTCCACACGCGGGACGCGGCGGAGCACTTGGACAGACCCCA 478

Db 3421 ACTCCGGCTGTCTCAAGGTGGTGAATCTCCACACGCGGGACGCGGCGGAGCACTTGGAGCA 3480

QY 479 ACGCGCTGTGGACACGCGGGACACGCGCGGGCAGGTGCGAACCTTATGGACAGACCCCA 538

Db 3481 ACGCGCTGTGGACACGCGGGACACGCGCGGGCAGGTGCGAACCTTATGGAGCACTTATGGACAGACCCCA 3540

QY 539 GGAACATTTGGCTGGAAGGACTACAGGCTTATAGTGGCACCTGACTCAGAGGCCCAAGA 598

Db 3541 GGAACATTTGGCTGGAAGGACTACAGGCTTATAGTGGCACCTGACTCAGAGGCCCAAGA 3600

QY 599 CCGGCTACATCAGAGTCTTAGTGCATGAAGGAAACAGGTGATGGCAGACTCAGGACCTTA 658

Db 3601 CTTGGCTACATCAGAGTCTTAGTGCATGAAGGAAACAGGTGATGGCAGACTCAGGACCTTA 3660

QY 659 TCTATGACCAAACTACGCTGCGCGGGCTGGGTCTATTTGTCTTCTCAAGAAATGG 718

Db 3661 TCTATGACCAAACTACGCTGCGCGGGCTGGGTCTATTTGTCTTCTCAAGAAATGG 3720

QY 719 TCTATTTCTCAGACCTCAAGTACGAATGACAGATATTTTAAACAAGATTTTCTGCATTTTC 778

Db 3721 TCATTTCTCAGACCTCAGTACGATGCGAGATATTTAAACAGATTTCTGCTCATTT 3780
Qy 779 CGGCAATGCCCTGTGCGATGCCATGGTCCCTAGACACCTCAGTTCAATGTGTGCTTGGG 838
Db 3781 CGGCAATGCCCTGTGCGATGCCATGGTCCCTAGACACCTCAGTTCAATGTGTGCTTGGG 3840
Qy 839 CTTCTCTCTAGCAGACCTCCTGCTCCCTGACCTTAACCTAGTCTGATGCTTCTACCTCC 898
Db 3841 CTTCTCTCTAGCAGACCTCCTGCTCCCTGACCTTAACCTAGTCTGATGCTTCTACCTCC 3900
Qy 899 TGCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATCAATGGAACCTCAGAGATG 958
Db 3901 TGCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATCAATGGAACCTCAGAGATG 3960
Qy 959 AACATCTAACCCCTAGAGGAAACAGTTTGGTGTATATAGAGACTTTATGTGGAGTAA 1018
Db 3961 AACATCTAACCCCTAGAGGAAACAGTTTGGTGTATATAGAGACTTTATGTGGAGTAA 4020
Qy 1019 AATTGGGCATGCATTACATTTCTTTCTTTGTTTAAAGAAATGAGCTTTACATA 1078
Db 4021 AATTGGGCATGCATTACATTTCTTTCTTTGTTTAAAGAAATGAGCTTTACATA 4080
Qy 1079 TAAATGTAAATACCTATTGTATTTATGTATATGAGTTGAAGGAAATACCTGTCATA 1138
Db 4081 TAAATGTAAATACCTATTGTATTTATGTATATGAGTTGAAGGAAATACCTGTCATA 4140
Qy 1139 AGCCATATGATAAATTAAGCATGAAGAAATATGCTGTAACCTTTTGGTCTTAAAGTT 1198
Db 4141 AGCCATATGATAAATTAAGCATGAAGAAATATGCTGTAACCTTTTGGTCTTAAAGTT 4200
Qy 1199 GTCACTATTCTGAATAGAGTTGCTCTACATGACACACAATCCCGCTAAATAATTA 1258
Db 4201 GTCACTATTCTGAATAGAGTTGCTCTACATGACACACAATCCCGCTAAATAATTA 4260
Qy 1259 TAAACAAGGGTCAATTCAAATTTGAAGTAAATTTTGTAGTGAAGGAGATTTAGAACAAC 1318
Db 4261 TAAACAAGGGTCAATTCAAATTTGAAGTAAATTTTGTAGTGAAGGAGATTTAGAACAAC 4320
Qy 1319 AGGCATAGCAATGACATAGCTACGATTAACCTAATCGAATGATTAACAGTTACAA 1378
Db 4321 AGGCATAGCAATGACATAGCTACGATTAACCTAATCGAATGATTAACAGTTACAA 4380
Qy 1379 AAATAAAGCACT 1438
Db 4381 AAATAAAGCACT 4440
Qy 1439 TTTCTATCAAGAAACAAACATCTCTGCAATGGGTGTGACGCGGTTCAGATGTGATTTG 1498
Db 4441 TTTCTATCAAGAAACAAACATCTCTGCAATGGGTGTGACGCGGTTCAGATGTGATTTG 4500
Qy 1499 GCAAAACCTCATTTAAGTAAAGGTTAGCAGAGCAAGTGGTGTCTTTAGTGTGCTTT 1558
Db 4501 GCAAAACCTCATTTAAGTAAAGGTTAGCAGAGCAAGTGGTGTCTTTAGTGTGCTTT 4560
Qy 1559 GTGCCGCTGTGGCGTGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCCG 1618
Db 4561 GTGCCGCTGTGGCGTGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCCG 4620
Qy 1619 AGAGAAACAGAGCAGACGCAAGGCGGCTTAAAGGCGCATCTAAACGCTATCTAGGCTTT 1678
Db 4621 AGAGAAACAGAGCAGACGCAAGGCGGCTTAAAGGCGCATCTAAACGCTATCTAGGCTTT 4680
Qy 1679 GGTAACTCGGACAGTGTCTTTTACCTGATTTGATGATACATTTTCAATTAAGCTTCCAGT 1738
Db 4681 GGTAACTCGGACAGTGTCTTTTACCTGATTTGATGATACATTTTCAATTAAGCTTCCAGT 4740
Qy 1739 TATAAATATTTTCTTAATATTTTAAAGTCACTATATAGAGTGAACCTTCCATTTACCAAGTAA 1798
Db 4741 TATAAATATTTTCTTAATATTTTAAAGTCACTATATAGAGTGAACCTTCCATTTACCAAGTAA 4800
Qy 1799 CTTATTTTAAATATGCTAGTAAACATATGATGATATATTTCTAGAAACAAACATCTAA 1858

Db 4801 CTTATTTTAAATATGCTAGTAAACATATGATGATATATTTCTAGAAACAAACATCTAA 4860
Qy 1859 TAAGTATATATCTGTGAAAATATGAGGCTTGATAATATTTAGGTTGTGACGATGAAGCA 1918
Db 4861 TAAGTATATATCTGTGAAAATATGAGGCTTGATAATATTTAGGTTGTGACGATG-AGCA 4919
Qy 1919 TGCTAGAAGCTGTAAACAGAAATACATAGAGAAATATAGGAGTTTATGATGAA-CTTTAA 1977
Db 4920 TGCT-GAAGCTGT-ACCGAAATACATAGAGAAATATGAGAGTGTATGATGAACCTTTAA 4977
Qy 1978 TATATATATTTGCGAGGATTTTATGTTCAATATTTGTTACTGTTTATCTATCTGCTGATA 2037
Db 4978 TATATATATGTCGCCAGGATTTTATGTTCAATATTTGTTACTGTTTATCTATCTGCTGATA 5037
Qy 2038 TGGAAATCTTTTAAATCAAAACGCTGAAAACGAAATCAGCAATTTAGTCTTTGCCAGGCACACC 2097
Db 5038 TGGAAATCTTTTAAATCAAAACGCTGAAAACGAAATCAGCAATTTAGTCTTTGCCAGGCACACC 5097
Qy 2098 CAATAACAGTCATGTGTATATGCAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 2157
Db 5098 CAATAACAGTCATGTGTATATGCAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 5157
Qy 2158 TGGTTTCTTTTCTTTTAAAGTTCATGATCTTCTGAGGAAATAGTCACTCATCCACATCC 2217
Db 5158 TGGTTTCTTTTCTTTTAAAGTTCATGATCTTCTGAGGAAATAGTCACTCATCCACATCC 5217
Qy 2218 ACATAAGGGTTTTAAAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 2277
Db 5218 ACATAAGGGTTTTAAAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 5277
Qy 2278 TTCCCTTTCTTTTAAATAGTCACTATTTCTATGCGAAACAGGACCATCATCACTTTT 2337
Db 5278 TTCCCTTTCTTTTAAATAGTCACTATTTCTATGCGAAACAGGACCATCATCACTTTT 5337
Qy 2338 AGTCTTATGATACATTTGATGATAAATTAATTTGTTGTTTCTTCTTCTTCTTCTTCTTCT 2397
Db 5338 AGTCTTATGATACATTTGATGATAAATTAATTTGTTGTTTCTTCTTCTTCTTCTTCTTCT 5397
Qy 2398 TCGTTGTGTTTGTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2456
Db 5398 TCGTTGTGTTTGTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5457
Qy 2457 CAACGAAGGTTGGGATACCTTAAATGATGAGCACTGTCAACAGCG 2504
Db 5458 CAACGAAGGTTGGGATACCTTAAATGATGAGCACTGTCAACAGCG 5505

RESULT 12

G06722
LOCUS human STS WI-7830, 2026 bp DNA linear STS 19-OCT-1995
DEFINITION G06722
ACCESSION G06722.1 GI:859967
VERSION STS; STS sequence; primer; sequence tagged site.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2026)
AUTHORS Hudson,T.
TITLE Whitehead Institute/MIT Center for Genome Research; Physically
MAPPED ESTs
JOURNAL Unpublished (1995)
COMMENT

Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: CGAGAAATAGTCACTCATCCC

Primer B: TTGACCCCAACACAGAAA
STS size: 347
PCR Profile:

Presoak:

Denaturation:

Annealing: 56 degrees C

Polymerization:

PCR Cycles: 35

Thermal Cycler:

Protocol:

Template: 10 na

Primer: each 5 pM

dNTPs: each 4 mM

Taq Polymerase: 0.025 units/ μ l

Total Vol: 20 ul

Buffer:

BUFFER:
MgCl2: 1 5 mM

MgCl₂: 1.3
KCl: 50 mM

NaCl: 50 mM
Tris-HCl: 10 mM

CH-93

Prepared with primer pairs derived from L12350 -- Unigene

FEATURES	SOURCE
1. Feature 1	Source 1
2. Feature 2	Source 2
3. Feature 3	Source 3
4. Feature 4	Source 4
5. Feature 5	Source 5
6. Feature 6	Source 6
7. Feature 7	Source 7
8. Feature 8	Source 8
9. Feature 9	Source 9
10. Feature 10	Source 10

```
1. 2026
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="770_E: 778_D.7;
933_F.7; 725_A.11"
1432. .1778
1432. .1453
complement(1759. .1778)
```

```

STS
primer_bind
primer_bind
1432. .1778
1432. .1453
complement(1759. .1778)
333_F_1_125_A_11*

```

Query Match	71.3%;	Score 1985;	DB 11;	Length 2026;
Best Local Similarity	99.0%;	Pred. No. 0;		
Matches 1985;	Conservative 0;	Mismatches 41;	Indels 0;	Gaps 0;

QY	760	ACAAGATTTCGTCATTTCCGGCAATGCCCTGTGTCATGCCATGGTCCTAGACACACTCAG	819
DB	1	ACAAGATTTCGTCATTTCCGGCAATGCCCTGTGTCATGCCATGGTCCTAGACACACTCAG	60
QY	820	TTCAATTGTGTCCTTGGGGCTTCTCTCTAGCAGCACCTCCTGTGCTCCCTTGACCTTAACT	879
DB	61	TTCAATTGTGTCCTTGGGGCTTCTCTCTAGCAGCACCTCCTGTGCTCCCTTGACCTTAACT	120
QY	880	CTGATGGTTCTTCACTCTGCCAGCAACCCCAACCCAAAGTGCTTTCAGAGGATAAATA	939
DB	121	CTGATGGTTCTTCACTCTGCCAGCAACCCCAACCCAAAGTGCTTTCAGAGGATAAATA	180
QY	940	TCAATGGAACTCAGAGATGAACATCTAACCCACTAGAGGAAACAGTTTGGTGATATATG	999
DB	181	TCAATGGAACTCAGAGATGAACATCTAACCCACTAGAGGAAACAGTTTGGTGATATATG	240
QY	1000	AGACTTTATGTGAGTGAAATTTGGGCATGCCATTACATTTGCTTTTCTTGTTGTTTAA	1059
DB	241	AGACTTTATGTGAGTGAAATTTGGGCATGCCATTACATTTGCTTTTCTTGTTGTTTAA	300
QY	1060	AAAGAAATGAGTTTACATATAAAAATGTAAATTACTTATTGTATTATGTGTATATGAGTT	1119
DB	301	AAAGAAATGAGTTTACATATAAAAATGTAAATTACTTATTGTATTATGTGTATATGAGTT	360
QY	1120	GAAGGGAATACCTGTGCATAAGCCATTATGATAAATTAAGCATGAAAAATTTGCTGAACT	1179
DB	361	GAAGGGAATACCTGTGCATAAGCCATTATGATAAATTAAGCATGAAAAATTTGCTGAACT	420
QY	1180	ACTTTTGGTGCTTAAAGTTGTCACTATTCTTGAATTAGAGTTGCTCTACAATGACACACA	1239
DB	421	ACTTTTGGTGCTTAAAGTTGTCACTATTCTTGAATTAGAGTTGCTCTACAATGACACACA	480
QY	1240	AATCCCGCTAAAATAAATTTAAACAAGGTCAAATTCAAATTTCAAGTAAATGTTTAA	1299

Qy	1530	AGCAAAGTCGGGTGCTTTAGCTGCTGCTTGTGCGGTGTGGCGTCTCTGCC	1589
Db	28031		
		AGCAAAGTCGGGTGCTTTAGCTGCTGCTTGTGCGGTGTGGCGTCTCTGCC	27972
Qy	1590	TGAGCTTCCTTTCCCAAGCTTTGCTGCTCAGAGGACACAGACAGACGCAAGGCCGAA	1649
Db	27971		
		TGAGCTTCCTTTCCCAAGCTTTGCTGCTCAGAGGAAACACAGACAGACGCAAGGCCGAA	27912
Qy	1650	AAGCGCATCTAAACGGGTATCTAGGCTTTGGTAACTGCGGACAAGTGTGCTTTACTCTGAT	1709
Db	27911		
		AAGCGCATCTAAACGGGTATCTAGGCTTTGGTAACTGCGGACAAGTGTGCTTTACTCTGAT	27852
Qy	1710	TTGATGATACATTTCACTAAGGTTCCAGTTATAAATATTTTGGTAAATTTATTTAAGTGA	1769
Db	27851		
		TTGATGATACATTTCACTAAGGTTCCAGTTATAAATATTTTGGTAAATTTATTTAAGTGA	27792
Qy	1770	CTATAGAATGCAACTCCATTTACCACTTAATTTTAAATATGCTAGTAAACACATATG	1829
Db	27791		
		CTATAGAATGCAACTCCATTTACCACTTAATTTTAAATATGCTAGTAAACACATATG	27732
Qy	1830	TAGTATAAATTTCTAGAAACAAACATCTAATAAGTATATAAATCCTGTGAAAAATATGAGGCT	1889
Db	27731		
		TAGTATAAATTTCTAGAAACAAACATCTAATAAGTATATAAATCCTGTGAAAAATATGAGGCT	27672
Qy	1890	TGATAATATTAGGTTGTGACGATGAAGCATGCTAGAGCTGTAACAGAAATACATAGAGAA	1949
Db	27671		
		TGATAATATTAGGTTGTGACGATGAAGCATGCTAGAGCTGTAACAGAAATACATAGAGAA	27612
Qy	1950	TAATGAGGAGTTTATCATGGAACCTT - AATATATAATGTTGCCAGCGATTTTAGTTCAT	2008
Db	27611		
		TAATGAGGAGTTTATGATGGAACCTTAAATATATAATGTTGCCAGCGATTTTAGTTCAT	27552
Qy	2009	ATTTGTTACTGTTATCTATCTGCTGTATATGGAATTCCTTTTAATTCAAACGCTCAAAACG	2068
Db	27551		
		ATTTGTTACTGTTATCTATCTGCTGTATATGGAATTCCTTTTAATTCAAACGCTCAAAACG	27493
Qy	2069	AATCAGCATTTAGTCTTGCGAGGCACACCCAAATAATCAGTCATGTGTAATATGCACAAGT	2128
Db	27492		
		AATCAGCATTTAGTCTTGCGAGGCACACCCAAATAATCAGTCATGTGTAATATGCACAAGT	27433
Qy	2129	TGTGTTTTGTTTTGTTTTTTTTTTTTGTTGGTTG --- GTTTTTTTGCTTTAAGTTCGATGAT	2184
Db	27432		
		TGTGTTTTGTTTTGTTTTTTTTTTTTGTTGGTTGTTTTTTTGTCTTTAAGTTCGATGAT	27373
Qy	2185	CTTTCTGCAGGAAATAGTCACATCCCACTCCACATAAGGGGTTTAGTAAGAGAAAGTCT	2244
Db	27372		
		CTTTCTGCAGGAAATAGTCACATCCCACTCCACATAAGGGGTTTAGTAAGAGAAAGTCT	27313
Qy	2245	GTCTGTCTGATGATGATAGGGGGGAAATCTTTTCCCTTTCTGTGTAATAGTCAATCACA	2304
Db	27312		
		GTCTGTCTGATGATGATAGGGGGGAAATCTTTTCCCTTTCTGTGTAATAGTCAATCACA	27253
Qy	2305	TTTCTATGCAAAACAGGAACGATCCATAACTTTAGTCTTAATGTACACATTGCAATTTGA	2364
Db	27252		
		TTTCTATGCAAAACAGGAACGATCCATAACTTTAGTCTTAATGTACACATTGCAATTTGA	27193
Qy	2365	TAAAAATTAATTTGTGTTTTCTTTGAGGTTGATCGTTG --- TGTGTTTTTGTCTGCACTT	2421
Db	27192		
		TAAAAATTAATTTGTGTTTTCTTTGAGGTTGATCGTTGTTGTTGTTTGTCTGCACTT	27133
Qy	2422	TTTTACTTTTTGCGTGTGAGCTGTATTTCCGAGA - CAAACGAGCGTGTGGGATATCTTCAT	2480
Db	27132		
		TTTTACTTTTTGCGTGTGAGCTGTATTTCCGAGACCAACGAGCGTTGGGATACTTCAT	27073
Qy	2481	TAAATGTACGCACTCTCAACAGCGTGCAGGTTTCTGTTCTGTGTTGTGGGFTCAACCG	2540
Db	27072		
		TAAATGTACGCACTCTCAACAGCGTGCAGGTTTCTGTTCTGTGTTGTGGGFTCAACCG	27013
Qy	2541	TACAATGGTGTGGGAATACGATGATGTGAATATTTAGAATGTACCATATTTTTGTGAAA	2600
Db	27012		
		TACAATGGTGTGGGATGACGATGATGTAGATTTTAGAATGTACCATATTTTTGTGAAA	26953
Qy	2601	TTATTTATGTTTTTCTAAACAAATTTATCGTATAGGTTGTAGAAACGCTCATGTGTTTTGC	2660

Db	26952	TTATTATTGTTTTTCTTAACAAATTTATCGTATAGGTTGATGAACGTCATGTTGTC	26893
QY	2661	CAAAGACTGTAATATTTATTTATGTTTCACATGGTCAAAATTTCCACTGAACCCCT	2720
Db	26892	CAAAGACTGTAATATTTATTTATGTTTCACATGGTCAAAATTTCCACTGAACCCCT	26833
QY	2721	GCACCTTAGCTAGAACCTCATTTTTTAAAGATTAAACACAGGAATAAATTTGTAAAAAGGT	2780
Db	26832	GCACCTTAGCTAGAACCTCATTTTTTAAAGATTAAACACAGGAATAAATTTGTAAAAAGGT	26773
QY	2781	TTTCT 2785	
Db	26772	TTTCT 26768	
RESULT 14			
HUMTHRSPD			
LOCUS			
DEFINITION Human thrombospondin mRNA.			
ACCESSION M81339			
VERSION M81339.1 GI:339678			
KEYWORDS thrombospondin.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
AUTHORS LaBell,T.L., Milewicz,D.J., Distech,C.M. and Byers,P.H.			
TITLE Thrombospondin II: partial cDNA sequence, chromosome location, and expression of a second member of the thrombospondin gene family in humans			
JOURNAL Genomics 12 (3), 421-429 (1992)			
MEDLINE 92217961			
PubMed 1553694			
COMMENT			
FEATURES			
source			
1..2780			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/cell_type="fibroblast"			
/tissue_type="connective"			
/dev_stage="adult"			
ORIGIN			
Query Match 60.8%; Score 1694; DB 9; Length 2780;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GAGGACTTGGACGGTGATGGACGGGTGATATTTGTTAAAGATGATTTTGACATGACAC 60	
Db	1085	GAGGACTTGGACGGTGATGGACGGGTGATATTTGTTAAAGATGATTTTGACATGACAC 1144	
QY	61	ATCCAGATATTGATGATGTGTCTCCTGAAAACAATGCCATCAGTGAGACAGACTTCAGG 120	
Db	1145	ATCCAGATATTGATGATGTGTCTCCTGAAAACAATGCCATCAGTGAGACAGACTTCAGG 1204	
QY	121	AACCTTCAGATGGTCCCTTGGATCCAAAGAGGACACCCAAATTGATCCCAACTGGGTC 180	
Db	1205	AACCTTCAGATGGTCCCTTGGATCCAAAGAGGACACCCAAATTGATCCCAACTGGGTC 1264	
QY	181	ATTCGCCATCAAGGCAAGGAGCTGTTTCAGACAGACCACTCGGACCCGGCATCGTGTA 240	
Db	1265	ATTCGCCATCAAGGCAAGGAGCTGTTTCAGACAGACCACTCGGACCCGGCATCGTGTA 1324	
QY	241	GGTTTTGACGAGTTTGGGTCCTGTGGACTTTCAGTGGCAATTTCTAGTAAACACTGACCGG 300	
Db	1325	GGTTTTGACGAGTTTGGGTCCTGTGGACTTTCAGTGGCAATTTCTAGTAAACACTGACCGG 1384	
QY	301	GACGACGACTATGCTGGCTTCGTCTTTGGTACCAAGTCAAGCAGCGGCTTCATGTGGTG 360	
Db	1385	GACGACGACTATGCTGGCTTCGTCTTTGGTACCAAGTCAAGCAGCGGCTTCATGTGGTG 1444	

QY	361	ATGTGGAAGCAGGTGACGAGACCTACTCTGGAGGACCAGCCACGCGGGCCCTATGGCTAC	420
DB	1445	ATGTGGAAGCAGGTGACGAGACCTACTCTGGAGGACCAGCCACGCGGGCCCTATGGCTAC	1504
QY	421	TCGGGGTGTCCTCTCAAGGTGTGAACTCGACACGCGGGACGGGCGAGCACTTGAGGAAC	480
DB	1505	TCGGGGTGTCCTCTCAAGGTGTGAACTCGACACGCGGGACGGGCGAGCACTTGAGGAAC	1564
QY	481	GCCTGTGGCACACGGGGAACGCGGGGCGAGTGCAGACCTTATGGCAGCACCCACG	540
DB	1565	GCCTGTGGCACACGGGGAACGCGGGGCGAGTGCAGACCTTATGGCAGCACCCACG	1624
QY	541	AACATTGGCTGGAAGGACTACACGGCCCTATAGTGTGGCACTGACTCAGAGGCCAAGACC	600
DB	1625	AACATTGGCTGGAAGGACTACACGGCCCTATAGTGTGGCACTGACTCAGAGGCCAAGACC	1684
QY	601	GGCTACATCAGAGTCTTAGTGCATGAAGGAACACGGTCAATGCAGACCTCAGACCTATC	660
DB	1685	GGCTACATCAGAGTCTTAGTGCATGAAGGAACACGGTCAATGCAGACCTCAGACCTATC	1744
QY	661	TATGACCAAACTACCGCTGGCGGGCGGCTGGGTCTATTGTCTCTCTCAAGAAATGGTC	720
DB	1745	TATGACCAAACTACCGCTGGCGGGCGGCTGGGTCTATTGTCTCTCTCAAGAAATGGTC	1804
QY	721	TATTTCTCAGACCTCAAGTACGAATGACAGAGATATTTAAACAAGATTGTGCTGCAATTCG	780
DB	1805	TATTTCTCAGACCTCAAGTACGAATGACAGAGATATTTAAACAAGATTGTGCTGCAATTCG	1864
QY	781	GCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTCATTTGGTGGTCCCTCGGCT	840
DB	1865	GCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTCATTTGGTGGTCCCTCGGCT	1924
QY	841	TCTCTCTTAGCAGCACCTCGTGCCTTGAACCTTAACCTCTGATGGTTCCTACCTCCTG	900
DB	1925	TCTCTCTTAGCAGCACCTCGTGCCTTGAACCTTAACCTCTGATGGTTCCTACCTCCTG	1984
QY	901	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAACCTCAGAGTAA	960
DB	1985	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAACCTCAGAGTAA	2044
QY	961	CATCTAACCCACTAGAGAAACCAAGTTTGGTGTATATGAGACTTTATGTGGAGTGAAAA	1020
DB	2045	CATCTAACCCACTAGAGAAACCAAGTTTGGTGTATATGAGACTTTATGTGGAGTGAAAA	2104
QY	1021	TTGGGCATGCCAATTACATTCGTTTTCTGTGTTCTTTTAAAAAGAAATCAAGTTTACATATA	1080
DB	2105	TTGGGCATGCCAATTACATTCGTTTTCTGTGTTCTTTTAAAAAGAAATCAAGTTTACATATA	2164
QY	1081	AAATGTAAATTAATTTATTTATTTATGTTATATCGAGTTGAGGGAATATCTGTGCATAAG	1140
DB	2165	AAATGTAAATTAATTTATTTATTTATGTTATATCGAGTTGAGGGAATATCTGTGCATAAG	2224
QY	1141	CCATTATGATAAAATTAAGCATGAAAAATTAATGCTGAACTACTTTTGGTGGCTTAAAGTTGT	1200
DB	2225	CCATTATGATAAAATTAAGCATGAAAAATTAATGCTGAACTACTTTTGGTGGCTTAAAGTTGT	2284
QY	1201	CATTAATCTGAAATTAGAGTTGCTCTACAATGCACACAAATCCGCTAAATAAATTAATA	1260
DB	2285	CATTAATCTGAAATTAGAGTTGCTCTACAATGCACACAAATCCGCTAAATAAATTAATA	2344
QY	1261	AACAGGGTCAATTCAAATTTGAAGTAATGTTTTAGTAAAGAGAGATTAGAAAGACAACAG	1320
DB	2345	AACAGGGTCAATTCAAATTTGAAGTAATGTTTTAGTAAAGAGAGATTAGAAAGACAACAG	2404
QY	1321	GCATAGCAAAATGACATAAGCTACCGATTAACTAATTCGGAAACATGTAAACAGTTTACAAA	1380
DB	2405	GCATAGCAAAATGACATAAGCTACCGATTAACTAATTCGGAAACATGTAAACAGTTTACAAA	2464
QY	1381	ATTAACGAACCTCTCTCTTGCTCAATCAATGAAGCCCTCATGTGCAGTAGAGATGCGAGTT	1440
DB	2465	ATTAACGAACCTCTCTCTTGCTCAATCAATGAAGCCCTCATGTGCAGTAGAGATGCGAGTT	2524

Qy	1441	TCATCAAGACGACAAACATCCTCTGCAAAATGGGTCTGACGGGTTCCAGATCTGGAATTTGGC	1500
Db	2525	TCATCAAGACGACAAACATCCTCTGCAAAATGGGTCTGACGGGTTCCAGATCTGGAATTTGGC	2584
Qy	1501	AAAACCTCATTTAAAGTAAAGGTTAGCAGACGAAAGTGGCGTCTTTAGCTGCTCTTGT	1560
Db	2585	AAAACCTCATTTAAAGTAAAGGTTAGCAGACGAAAGTGGCGTCTTTAGCTGCTCTTGT	2644
Qy	1561	GCCGTTGTGGCGTGGGGAGGCTCTCTGCTGAGACTTCTTCCCAGCTTTGCTGCTGCTGAG	1620
Db	2645	GCCGTTGTGGCGTGGGGAGGCTCTCTGCTGAGACTTCTTCCCAGCTTTGCTGCTGCTGAG	2704
Qy	1621	AGGAACAGAGCAGACGACAGCGCCGAAAGCGCATCTTAACGCGTATCTAGGCTTTGG	1680
Db	2705	AGGAACAGAGCAGACGACAGCGCCGAAAGCGCATCTTAACGCGTATCTAGGCTTTGG	2764
Qy	1681	TAACTGCGGACAAG	1694
Db	2765	TAACTGCGGACAAG	2778
RESULT 15			
BV177085/c			
LOCUS			
DEFINITION	BV177085	921 bp DNA linear	STS 10-JUN-2004
	scnm92668 Human DNA (Sequenom)	Homo sapiens STS genomic, sequence	
ACCESSION			
VERSION	BV177085		
KEYWORDS	BV177085.1	GI:48013080	
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 921).		
TITLE	Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A. Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions		
JOURNAL	Genome Res. (2004) In press		
COMMENT			
	Contact: Andreas Braun		
	Pharmaceuticals division		
	Sequenom, Inc.		
	3595 John Hopkins Court, San Diego, CA 92121, USA		
	Tel: 18582025018		
	Fax: 18582023020		
	Email: abraun@sequenom.com		
	Primer A: No primer sequence submitted		
	Primer B: No primer sequence submitted		
	STS size: 921.		
FEATURES			
source	Location/Qualifiers		
	1...921		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/clone_lib="Human DNA (Sequenom)"		
	<1...>921		
STS			
ORIGIN			
	Query Match	29.0%;	Score 809; DB 11; Length 921;
	Best Local Similarity	98.2%;	Pred. No. 6.1e-168;
	Matches 869; Conservative	1; Mismatches 5; Indels 10; Gaps 5;	
Qy	1910	GATGAAGCATGCTAGAAGCTGTAAACAGAAATACATAGAGAAATAGAGAGGTTTATGATGG	1969
Db	915	GATGAAGCATGCTAGAAGCTGTAAACAGAAATACATAGAGAAATAGAGAGGTTTATGATGG	856
Qy	1970	AACCTT-AATATATATGTTGCCAGCAATTTAGTCAATATTTGTTACTGTTATCTATC	2028
Db	855	AACCTTAAATATATATGTTGCCAGCAATTTAGTCAATATTTGTTACTGTTATCTATC	796
Qy	2029	TGCTGTATGCAATCTTTTAATTCACACGCTGAAAACGAATCAGCAATTTAGTCTTCTGCC	2088

Db 795 TGCTGATATGGAATCTTTTAAATCAACGCTGA AAA-GAATCAGCAITTAGTCTTGCC 737
QY 2089 AGGCACACCCAAATAATCAGTCATGTGTAATATGACAAAGTTGTGTTTGTGTTTGTGTTT 2148
Db 736 AGGCACACCCAAATAATCAGTCATGTGTAATATGACAAAGTTGTGTTTGTGTTTGTGTTT 677
QY 2149 TTTGTTGGTTG---GTTTTTTTCTTTAAAGTTGCATGATCTTTCTGCAGAAATAGTCA 2204
Db 676 TTTGTTGGTTGGTTTGTGTTTGTGTTTAAAGTTGCATGATCTTTCTGCAGAAATAGTCA 617
QY 2205 CTCATCCCACTCCACATAAGGGTTTAGTAAGAGAAGTCTGTCTGTCTGATGATGGATAG 2264
Db 616 CTCATCCCACTCCACATAAGGGTTTAGTAAGAGAAGTCTGTCTGTCTGATGATGGATAG 557
QY 2265 GGGGCAATCTTTTCCCTTTCTGTTTAAAGTATCACATTTCTATGCCAAACAGAAC 2324
Db 556 GGGGCAATCTTTTCCCTTTCTGTTTAAAGTATCACATTTCTATGCCAAACAGAAC 497
QY 2325 GATCCATAACTTTAGTCTTAAAGTATCACATTTGCAATTTTGATAAAATTAATTTTGTGTTT 2384
Db 496 RATCCATAACTTTAGTCTTAAAGTATCACATTTGCAATTTTGATAAAATTAATTTTGTGTTT 437
QY 2385 CCTTTGAGGTGATCGTGG---TGTTGTTTGGTGCACATTTTACATTTTGGGTGGGA 2441
Db 436 CCTTTGAGGTGATCGTGGTTGTTGTTTGGTGCACATTTTACATTTTGGGTGGGA 377
QY 2442 GCTGTATTTCCCGAGA-CAACGAAGCGTTGGGATACCTTCAATTAATGTAGCGACTGTCAAC 2500
Db 376 GCTGTATTTCCCGAGACCAACGAAGCGTTGGGATACCTTCAATTAATGTAGNGACTGTCAAC 317
QY 2501 AGCGTGCAGGTTTCTGTTTCTGTTGTTGGGTCAACCGTACAATGGTGGGANTGAC 2560
Db 316 AGCGTGCAGGTTTCTGTTTCTGTTGTTGGGTCAACCGTACAATGGTGGGANTGAC 257
QY 2561 GATGATGCAATATTTAGAAATGTACCATATTTTGTAAATTAATTTATGTTTTCTAAAC 2620
Db 256 GATGATGCAATATTTAGAAATGTACCATATTTTGTAAATTAATTTATGTTTTCTAAAC 197
QY 2621 AAATTTATCGTATAGGTTGATGAAAGCTCATGTGTTTGGCCAAAGACTGTAAATATTTAT 2680
Db 196 AAATTTATCGTATAGGTTGATGAAAGCTCATGTGTTTGGCCAAAGACTGTAAATATTTAT 137
QY 2681 TTATGTGTTTCAATGTCAAAAATTTCCACCTGAAAAACCTGCACTTAGTAGAACCTCAT 2740
Db 136 TTATGTGTTTCAATGTCAAAAATTTCCACCTGAAAAACCTGCACTTAGTAGAACCTCAT 77
QY 2741 TTTTAAAGATTAAACACAGGAATAAATTGTA AAAAGGTTTTCT 2785
Db 76 TTTTAAAGATTAAACACAGGAATAAATTGTA AAAAGGTTTTCT 32

Search completed: November 23, 2004, 04:30:21
Job time : 11538.2 secs

This Page Blank (uspto)